Final Report for AOARD Grant 10-4102

Robust Multi-Agent Sensor Network Systems

8th May 2012

Principal Investigator: Brian D O Anderson
- e-mail address: brian.anderson "at" anu.edu.au
- Institution : National ICT Australia (NICTA)
- Mailing Address : National ICT Australia Limited, Locked Bag 8001, Canberra, ACT 2601, Australia
- Phone : +61 2 6125 8667
- Fax : +61 2 6125 8660 (likely to be ceasing operation)

Co- Investigators: (1) Adrian Bishop (adrian.bishop "at" nicta. com. au)
(2) Guoqiang Mao ("at" sydney .edu.au)
(3) Changbin (Brad) Yu (brad.yu "at" anu.edu.au)

Period of Performance; 06 May 2010 – 05 May 2012

Abstract: Results obtained with the grant fall into several groupings: (a) Combinatorial conditions on the graphical representation of a two-dimensional sensor network that will guarantee localizability of the network in the event of loss of any p sensors and/or q links in the network, for nonnegative integers p and q; (b) analysis of the effects of measurement error on the quality of localization of sensor positions in a sensor network, or more generally a target being localized; (c) the derivation of a measure, including algorithms for computing it, of the quality of connectivity of a network modeled by a graph with nodes and links, and in which the individual links are operative with defined a priori probabilities, and the probability that any one link is operative is independent of the probability that any other link is operative; (d) connectivity and capacity of networks with randomly positioned nodes and probabilistic channel models; (e) Doppler localization problems and miscellaneous multiagent problems

Introduction

Background to the problems considered

The applications context of the work is sensor networks in the first instance. A sensor network comprises a collection of sensors, some of which are at known positions due to *a priori* measurement or the equipping of such sensors with GPS, while others are at unknown positions. Sensors without positioning capability however can measure distances to other sensors provided these other sensors are sufficiently close ('within the sensing radius'). (Such measurements are distinct from those which the sensor network is meant to collect, which can relate to intruders, fire, water, chemicals, etc). *Sensor network localization* is the task of determining from all this data the positions of all the sensors. (Ability to do this is an obviously desirable feature if the collected information relating to intruders, fire, etc is to be of use.) A sensor network is termed *localizable* when such determination is possible.

Localizability is not the same concept as localization; localizability is a prerequisite property

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
10 MAY 2012	**Final**	**06-05-2010 to 05-05-2012**

4. TITLE AND SUBTITLE
Robust Multi-Agent Sensor Network Systems

5a. CONTRACT NUMBER
FA2386-10-1-4102

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Brian Anderson; Guoqiang Mao; Changbin Yu; Adrian Bishop

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NATIONAL ICT AUSTRALIA (NICTA) LIMITED ,LOCKED BAG 8001, CANBERRA ACT 2601 ,AUSTRALIA,AU,AU,2601

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AOARD, UNIT 45002, APO, AP, 96338-5002

10. SPONSOR/MONITOR'S ACRONYM(S)
AOARD

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AOARD-104102

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Results obtained with the grant fall into several groupings: (a) Combinatorial conditions on the graphical representation of a two dimensional sensor network that will guarantee localizability of the network in the event of loss of any p sensors and/or q links in the network, for nonnegative integers p and q; (b) analysis of the effects of measurement error on the quality of localization of sensor positions in a sensor network, or more generally a target being localized; (c) the derivation of a measure, including algorithms for computing it, of the quality of connectivity of a network modeled by a graph with nodes and links, and in which the individual links are operative with defined a priori probabilities, and the probability that any one link is operative is independent of the probability that any other link is operative; (d) connectivity and capacity of networks with randomly positioned nodes and probabilistic channel models; (e) Doppler localization problems and miscellaneous multi-agent problems.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**264**	

required prior to localization. The CI in conjunction with others developed a graphical (combinatorial) characterization of the localizability property that was published several years ago. Extensions of that work also included identification of classes of network for which localization could be easily achieved (ease of localization being characterized more or less by computational complexity of the localization algorithm).

There are shortcomings in these results, since they rest on idealized situations not necessarily met in practice. For example:

- Links or sensors themselves may fail, particularly in a large network. The existing theory characterizing network localizability made no allowance for this.
- Measurements of intersensor distances are noisy. The existing theory made no attempt to relate the errors in intersensor distance measurement to errors in the localized positions of those sensors lacking a priori position knowledge

The failure of links or sensors in a sensor network is a particular case of failure of links or nodes in a normal communications network. It can be the case that individual link quality in a network can be characterized by a simple probability, viz. the probability that the link will function properly on any occasion that it passes a message. Most message passing in a network however involves transmission of a message over a series of links, with multiple paths often existing between source and destination. A method has been lacking for assembling the individual link qualities (probabilities) into a measure for the overall quality of a network.

In many networks, whether sensor networks or ad hoc networks or mobile networks, the node positions are essentially random. For this reason and especially for large scale networks, it can be desirable to characterize the performance of such networks in terms of statistical parameters, such as node density, rather than via deterministic graphical properties. Accordingly, techniques are needed to characterize connectivity and throughput in these random situations.

In the course of our work, several quite specific applications problems presented themselves. These included the use of measurements of Doppler shift for localization. In a number of military situations, there can be a single high speed target, and cooperative *but spatially separated sensors*.

Problems considered

The problems considered in the research were driven by the mismatch between available theory and the demands of real-world application recorded above. The principal problems addressed were originally as follows:

- Finding combinatorial conditions on the graphical representation of a two-dimensional sensor network for the localizability property to be retained in the event of loss of any p sensors and/or q links in the network
- Explaining for a localizable sensor network how errors in inter-sensor distances translate into errors in sensor position estimates (and more generally, how sensing errors in any localization scenario affect the quality of position estimate of a target)

While the results being sought were primarily analytic in character (given a network, say what properties it has), it was our hope that the results would give *design insights* as well.

Starting in the first year, but continuing much more in the second year, additional problems were addressed, motivated by those already considered

- Defining a quality measure for a network (not a sensor network) characterized by a set of nodes and a set of links with each link functioning with a defined probability, and such that the functioning or otherwise of two links are independent events
- Examining connectivity and throughput properties of networks where nodes are randomly positioned (and may be moving) and channel models for transmission between nodes are probabilistic
- Localization using Doppler measurements and sundry other multiagent problems, involving a mixture of formations, sensor networks and consensus ideas.

Publications are recorded at the end, in groupings corresponding to the above five bullet points.

Contextual matters

The proposal provided that interaction with US workers would occur, and that presentations would be made at US conferences. Interactions did occur, particularly with Professor S Dasgupta (University of Iowa) and Professor A S Morse (Yale University), and presentations at US conferences did occur. See the publication list for coauthored papers, and US conference papers.

Much research was also linked with collaborative work with Defence Science and Technology Organisation, Australia (DSTO); the focus of that work was cooperative localization of targets.

Outline of main results

Combinatorial conditions on the graphical representation of sensor networks

A graphical representation of a sensor network is defined by a graph $G=(V,E)$ in the usual sense, where V is the vertex set (abstracting the sensors) and E is the edge set (abstracting those sensor pairs for which the intersensor distance is known). Sensors are divided into those whose absolute positions are known, call them anchor sensors, and ordinary sensors.

Localizability of a sensor network in an ambient two-dimensional space occurs if and only if three conditions simultaneously hold (these conditions are reviewed in the papers in question, which give the original references; among these original references are papers by the Principal Investigator)

- The graph of the network is 3-connected (k-connectedness for any positive integer k is a standard concept in graph theory; it means that between any two vertices, k nonintersecting paths, i.e. connected sequences of edges linking one vertex to the other, can be found)
- There are 3 or more noncollinear anchor nodes
- The graph of the network is *redundantly rigid.* (Rigidity is a graph theory concept, capable of formal definition naturally, which captures the notion that the graph is a pictorial representation of a rigid, as opposed to flexing, structure. For example, a graph corresponding to a quadrilateral with a diagonal is a rigid graph, while a graph corresponding to a quadrilateral without a diagonal is not rigid. 'Redundantly rigid' means, in structural terms, that there are extra structural members beyond those

required to ensure rigidity; more precisely, rigidity is retained when any one graph edge is removed.

Our key conclusions can be summed up as follows and are obtained in the papers listed under the heading '*Redundant Localizability'*.

- Conditions ensuring retention of localizability given sensor loss are more demanding than conditions ensuring retention of localizability given link loss: in fact, if a network is localizable after loss of any p sensors and q links, it will be localizable after loss of any $p+q$ links
- In general, we could not find identical necessary and sufficient conditions for retention of localizability after loss of p sensors or of q links, though the difference between necessary and sufficient conditions was not great.
- With small enough values of p and q, identical necessary and sufficient conditions could however be found. For example, a network with $|V|$ sensors including 3 or more noncollinear anchors and $2|V|$ links is localizable when any single nonanchor sensor or link is lost, given an easily checked well-distributedness property of the links. Designs are easy.
- Networks retaining localizability given loss of sensors and/or links necessarily are of higher density that a network which simply has to be localizable. Density can be measured by the ratio of links to sensors, or separately by the maximum value of k for which the network is k-connected. For example, to be tolerant of the loss of q links, it is necessary that the ratio of edges to vertices exceed $(1/2)(q+3)$.
- Algorithms for verifying retention of localizability given loss of p sensors and q links with general p,q, i.e. algorithms checking either the necessary or sufficient conditions, can be very complicated. However, there is one very easily checked sufficient condition: if a network is $(p+q+6)$-connected and there are $p + 3$ or anchors, no three of which are collinear, then it will be localizable after loss of p sensors and q links. We note too that a necessary connectivity condition is simply that the network be $(p+q+3)$-connected.
- A network whose sensors are randomly distributed according to say a Poisson distribution (and there are $p + 3$ or anchors, no three of which are collinear) will, with very high probability, be $(p+q+6)$-connected (and thus robust to loss of p sensors and q links) given a sufficiently large (and computable) density parameter for the underlying Poisson distribution.

The effect of measurement errors

The first issue we sought to address was to understand how errors in intersensor distances translate into errors in localized sensor positions. We first undertook this study for a power law, and found that, roughly speaking, for a one-dimensional network (such as might be used along power lines, or a road for example), errors in position estimates would grow as the square of the minimal hop count to an anchor. Work was then undertaken on two-dimensional sensor networks and recently a conference paper on that topic was published. The results are so far quite limited, but new techniques have been developed. We considered the scenario where a single sensor was localized using noisy range measurements from a number of anchors, which were randomly located. We found that even with a surprisingly small number of anchors, the error in the sensor position estimate was close to Gaussian; there is a sort of spatial central limit theorem applying, which says that if you average nongaussian random variables, the average is nearer to being Gaussian. [What remains to be done, and will be hard, is the extension of this result to estimating sensor positions with noisy measurements from other sensors that are not anchors.]

Other work under this heading recorded in the publications was as follows:

- For a network in which sensors are Poisson distributed with known density and sensing radius, we found a method of estimating an intersensor distance between any pair of sensors that could sense one another based on determining for two sensors the number of other sensors that were within the sensing radius of both, and the number of other sensors that were within the sensing radius of just one. This procedure leads to surprisingly accurate localization using the estimated distances
- Localizing a sensor network may be computationally very expensive. We found a procedure based on localizing subnetworks, such that if each subnetwork could be separately localized, then it was easy to localize the whole network but gluing together the individual pieces
- We obtained a general procedure for approximation computation of bias—which is nothing more than a systematic error—in localization problems. (In work with DSTO, the method was validated on trial data)

Finding a network quality measure

We consider a graphical model of a network with the vertices corresponding to network nodes and the edges of the graph corresponding to the network links. Nodes fall into two classes: those that are internal to the network, and those that are source/destination nodes (if a node can be a source node, it can also be a destination node, and vice versa). Associated with each edge is a weight, corresponding to the probability that the link is operative. Link failures are independent of each other. While there are many matrices that can be associated with a graph, including a weighted graph, that reflect aspects of connectivity, e.g. adjacency matrix, Laplacian matrix, etc, none of them is useful here.

Our main results are as follows:

- We have defined a *probabilistic connectivity matrix,* with the property that the *ij* entry measures the probability that there is a path from node *i* to node *j* (Here, the nodes in question must be source/destination nodes)
- We obtained a (nontrivial) algorithm to compute this matrix
- The matrix is symmetric, nonnegative definite and has nonnegative entries; its largest eigenvalue (Perron eigenvalue) is a single measure of the network quality
- Preliminary results suggest experiments can be conducted at the source/destination nodes to determine the largest eigenvalue when the network's internal structure or link probabilities are unknown
- The matrix yields various topological properties of a network:
    - How reliably a node can communicate with another node
    - Existence of any single point-of-failure (critical nodes)
    - Other eigenvalues and their multiplicity say something about the topology, e.g. the number of strongly connected components

Random Network Problems

We considered networks where nodes are randomly placed in a given region. In some of our studies, these nodes are considered to be stationary while in others, these nodes are considered to be mobile. A link exists between a pair of nodes if their Euclidean distance is smaller than a given threshold, known as the transmission range. Under the above setting, we studied the connectivity property and the information propagation process of these networks under a series of scenarios not yet addressed in the literature.

Our results are as follows:

- We have found the necessary and sufficient condition for a network with stationary nodes to be connected where the link between a pair of nodes is subject to the detrimental impact of interference caused by other simultaneously active transmissions.
- We derived analytical expression for the capacity of the above network.
- Considering a vehicular network with vehicular nodes deployed along a line and with regularly placed roadside infrastructure, we found the probability that a randomly chosen node can access the infrastructure either directly or via other nodes as relays, and the probability that all nodes can access at least one infrastructure node. The two probabilities are important measures of network performance from a user and a network service provider perspective respectively.
- Considering a mobile network in two-dimensional space where each node is initially randomly placed and also moves randomly in a confined region, we found the information propagation speed of the network and the percentage of nodes that can receive the information of a randomly chosen node.

Doppler Localization Problems and Miscellaneous Multiagent problems

The main question we studied was how many spatially separated sensors for Doppler shift would be required to localize a fast moving target. The main conclusions we found were as follows:

- Four spatially separated sensors at generic positions and measuring Doppler shift alone can localize the target (and determine its velocity) up to a finite number of possibilities; disambiguation may be possible depending on *a priori* knowledge
- With five spatially separate sensors at generic positions, there is no ambiguity, and with six spatially separated sensors, arbitrary positions are satisfactory, except that sensors cannot be collocated.
    - One aspect of this work showed for the first time that although five sensors are generically sufficient for position and velocity estimation there are non-trivial counterexamples where five sensors still lead to ambiguity in the both the position and velocity estimate. This result stems solely from the nonlinear nature of the equations describing the Doppler shifts
- With noisy measurements, there are good and bad geometries, and of course fewer Doppler measurements will be enough if other measurements such as range or bearing should be available.

Other miscellaneous work included interaction with Professor A S Morse of Yale University on problems of consensus, which can have application to rendezvous problems, flocking problems (where a group of agents has to acquire a common velocity,) and estimation problems in a large scale system where agents must share noisy estimates of some underlying physical variable to arrive at a common estimate.

Triggered by our interest in retaining functionality of a sensor network in the even of sensor or link failure, we also started to look at security problems in which malicious interference with a network operation was a possibility.

Future Work Possibilities

The work has opened up a number of fronts on which further research can be pursued. We list some here;

- Networked control and estimation underpins the core technology in numerous critical infrastructure systems; e.g. the electricity, transport and many defence systems. However, the networked nature of modern control systems often make them vulnerable to a deliberate attack; e.g. many industrial control systems may even be accessible through the internet. As an outgrowth from our work on ensuring redundancy in sensor networks so that they retain functionality given node or link failure, one can seek to consider design methods to secure redundancy giving protection against attacks. The broad question is: How can networked control and estimation systems be protected from malicious cyber attackers? An emphasis on robust control, estimation and robust security is a key driver of this avenue of research that is already being considered within NICTA and the ANU. The outcome of this work is envisioned to be a suite of algorithms, tools and design considerations for safeguarding networked, industrial, control and estimation systems from attack.
- As the examples above cited above imply, many modern networked systems contain a mixture of control systems and communication systems. Such systems also are never fixed; devices join, links are made and broken etc. It is clear that the design of the communication system part of these complex systems should not be separated from the design of the control system part. Indeed, a seamlessly integrated system can lead to an even greater benefit and in many cases is a prerequisite for the entire system to perform satisfactorily. Our goal would be to develop tools for simultaneous communication network design and control system design to cater for the intricate demands of these complex systems and to integrate and exploit the characteristics and dynamics of complex systems.

Milestones

The following milestones were agreed for the second year.

Item	Status
Submit two papers on formations and/or sensor networks retaining *global* rigidity and localizability properties in event of loss of p nodes (and/or q communication links)	Done
Submit two papers on the connectivity of wireless sensor networks where sensors are connected following some realistic channel model and connections are subject to interference	Done
Submit paper on at least one other problem (e.g. distinguishing worst case and average case of robustness, error and bias analysis, deterministic gossip problem, Doppler-based	Easily exceeded

estimation etc)	
Arrange visit by a second US visitor and provide a report on expected downstream outcome	Done; Prof AS Morse visited in March 2012
Attend at least one international conference in the US to disseminate outcomes and networking	Done. PI and three CIs all did this.
Submit final report [both years] to AOARD	This document is the final report

Publications

Redundant Localizability

S. A. Motevallian, C. Yu and B.D.O Anderson, Robustness to the Loss of Multiple Nodes in the Localizability of Sensor Networks, 18th World Congress of the International Federation of Automatic Control (IFAC 2011), pages 7836–7841, Milan, Italy, August-September 2011.

S. A. Motevallian, C. Yu and B.D.O Anderson, Multi-Agent Rigid Formations: A Study of Robustness to the Loss of Multiple agents, Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC 2011), pages 3602-3607, Orlando, USA, December 2011.

C. Yu, S. Dasgupta and B.D.O. Anderson, Network Localizability with Link or Node Losses, Proc. of the 49th IEEE Conference on Decision and Control, pp. 402-407, Atlanta, USA, Dec 2010.

C. Yu, S. Dasgupta and B.D.O. Anderson, Redundant Localizability of Sensor networks, submitted to Automatica.

Other Aspects of Localization

B. Huang, C. Yu and B.D.O. Anderson, Analyzing Localization Errors in One-Dimensional Sensor Networks, Signal Processing, Vol. 92, no. 2,, 2012, pp. 427-438.

B. Huang, T. Li, B.D.O. Anderson and C. Yu, On the Performance Limit of Sensor Localization, Proc. of the 50th IEEE Conference on Decision and Control and European Control Conference, Orlando FL, Dec 2011.

B. Huang, C. Yu, B.D.O. Anderson and G. Mao, Estimating Distances via Connectivity in Wireless Sensor Networks, Wireless Communications and Mobile Computing, 2012, accepted for publication

B. Huang, C. Yu, B.D.O. Anderson and G. Mao, Connectivity-based Distance Estimation in Wireless Sensor Networks, Proc. of the IEEE GLOBECOM, pp 1-6, Miami, Florida, Dec 2010.

S. A. Motevallian, L. Xia and B.D.O Anderson, Localization of Sensor Networks using Bilateration and Four-Bar Linkage Mechanisms, Submitted to IEEE GlobeCom

12.

Y. Ji, C. Yu, B.D.O. Anderson and S.P. Drake, A Generic Bias-Correction Method with Application to Scan-Based Localization, Proc. of IEEE ICCA, Chile, Dec 2011.

C. Yu, B. Huang, H. Chee and B.D.O. Anderson, Noisy Localization on the Sphere, International Journal of Intelligent Defence Support System, Vol. 4, no. 4, 2011, pp. 328-350.

Quality of Network Connectivity

S. Dasgupta and G. Mao, On the Quality of Wireless Network Connectivity, submitted to IEEE Globecom 2012.

G. Mao, Wireless Multihop Networks: Current Research and Future Challenges, accepted to appear in Journal of Communications, Special Issue on Future Directions in Computing and Networking, (an extended version of the invited position paper published in International Conference on Computing, Networking and Communications 2012).

T. Yang, G. Mao and W. Zhang, Connectivity of Large-Scale CSMA Networks, accepted to appear in IEEE Transactions on Wireless Communications.

Z. Zhang, G. Mao and B. Anderson, On the Hop Count Statistics in Wireless Multi-hop Networks Subject to Fading, accepted to appear in IEEE Transactions on Parallel and Distributed Systems.

Z. Zhang, G. Mao and B.D.O. Anderson, On the Information Propagation Process in Mobile Vehicular Ad Hoc Networks, IEEE Transactions on Vehicular Technology, Vol. 60, No. 5, pp. 2314-2325, 2011

S.C. Ng, W. Zhang, Y. Yang and G. Mao, Analysis of Access and Connectivity Probabilities in Vehicular Relay Networks, IEEE Journal on Selected Areas in Communications--Special Issue Vehicular Communications and Networks, Vol. 29, No. 1, pp. 140-150, 2011

T. Yang, G. Mao and W. Zhang, Capacity of Interference-limited Three Dimensional CSMA Networks, accepted to appear in IEEE ICC 2012.

Z. Zhang, G. Mao and B.D.O. Anderson, On the Information Propagation Process in Multi-lane Vehicular Ad-hoc Networks, accepted to appear in IEEE ICC 2012

Z. Zhang, G. Mao and B.D.O Anderson, “On Information Dissemination in Infrastructure-based Mobile Ad-hoc Networks, accepted to appear in IEEE WCNC 2012.

Doppler Localization Problems and Miscellaneous Multiagent problems

A.N. Bishop and M. Smith, Remarks on the Cramer-Rao Inequality for Doppler-Based Target Parameter Estimation, In proceedings of the 6th International Conference on Intelligent Sensors, Sensor Networks and Information Processing

(ISSNIP 2010), Brisbane, Australia, December 2010.

I.S. Shames, A.N. Bishop, M. Smith and B.D.O. Anderson, Analysis of Target Velocity and Position Estimation via Doppler-Shift Measurements, In proceedings of the 2011 Australian Control Conference (AUCC), Melbourne, Australia, November 2011.

I.S. Shames, A.N. Bishop, M. Smith and B.D.O. Anderson, Target Velocity and Position Estimation via Doppler Shift Measurements, to appear in the IEEE Transactions on Aerospace and Electronic Systems, 2012.

A.N. Bishop, A Robust Reachability Review for Control System Security, In proceedings of the 2011 Australian Control Conference (AUCC), Melbourne, Australia, November 2011.

J. Liu, S. Mou, A.S. Morse, B.D.O. Anderson and C. Yu, Deterministic Gossiping, Proceedings of the IEEE, vol. 99, no. 9, pp. 1505-1524, 2011.

J. Liu, A.S. Morse, B.D.O. Anderson and C. Yu, Contractions for Consensus Processes, Proc. of the 50th IEEE Conference on Decision and Control and European Control Conference, Orlando FL, Dec 2011.

J. Liu, S. Mou, A.S. Morse, B.D.O. Anderson and C. Yu, Request-Based Gossiping, Proc. of the 50th IEEE Conference on Decision and Control and European Control Conference, Orlando FL, Dec 2011.

Attachments: Publications listed above.

DD882: Included separately

2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Multi-Agent Rigid Formations: A Study of Robustness to the Loss of Multiple agents

S. Alireza Motevallian, Changbin Yu, Brian D.O. Anderson

***Abstract*—In this paper we study the robustness of information architectures to control a formation of autonomous agents. If agents are expected to work in hazardous environments like battle-fields, the formations are prone to multiple agent/link loss. Due to the higher severity of agent loss than link loss, the main contribution of this paper is to propose information architectures for shape-controlled multi-agent formations, which are robust against the loss of multiple agents. A formation is said to be rigid if by actively maintaining a designated set of inter-agent distances, the formation preserves its shape. We will use the *rigidity* theory to formalize the robust architecture problem. In particular we study the properties of formation graphs which remain rigid after the loss of any set of up to $k-1$ vertices. Such a graph is called k-vertex rigid. We provide a set of distinct necessary and sufficient conditions for these graphs. We then show that 3-vertex rigidity is the highest possible robustness one can achieve by just adding a small number of edges to a minimally rigid graph, i.e. retention of rigidity given the loss of 3 or more agents of a formation requires many more inter-agent distances to be specified than when maintaining rigidity with no, one or two agent losses. Based on this result, we further focus on 3-vertex rigid graphs and characterize a class of information architectures (with minimum number of control links) which are robust against the loss of up to two agents.**

***Index Terms*—Formation Control, Robustness, Rigidity, Redundant Rigidity**

I. Introduction

Recently, autonomous agents and specifically UAVs (unmanned aerial vehicles) have found significant interests among researchers. UAVs have become an enabling technology in military application such as surveillance and reconnaissance over several decades [4], [12]. Today, there is also an increasing interest in UAVs for civil application such as environmental monitoring and exploration [1], [7]. In many applications it is desirable to have several autonomous UAVs flying in a formation [7]. The *formation control task* is to control, normally in a distributed manner, a set of inter-agent distances such that a prescribed shape is achieved and the formation moves as a cohesive whole [1], [5]. The reason is that when these agents, engaged in surveillance or exploration missions, move in a formation with a specific shape, they usually synthesize an antenna of magnitude far larger than a single agent [1], [14]. For certain distributed antenna shapes, this can lead to better sensitivity in target detection or localization over the area of interest.

Many of these missions for UAVs enforce the existence of a high level of autonomy of the vehicles, specifically in hazardous environments or in the presence of communication blackouts [9]. This autonomy can provide higher precision, fault tolerance and reliability in the mission accomplishment.

In hazardous environments such as battle-fields, the robustness of the formation is of high importance [9], [4]. Robustness here, means ensuring a successful accomplishment of the mission in different operational scenarios, when both some agents and communication links may be lost due to mechanical failure, enemy attack, jamming, an agent leaves the formation deliberately, etc.

In [4], as cited in [17], some vulnerability issues of UAVs are studied based on the actual records of the past and the potential issues are mainly classified as

- communication link loss as a result of jamming or occlusion
- enemy attack of one or more UAVs in the formation
- loss of an agent or communication link as a result of a mechanical or electrical failure, without enemy attack but possibly due to environmental changes (heat, wind, etc.).

Based on the above study the authors concluded the necessity of reducing the vulnerability of UAV to such threats.

It is evident that the robustness of the formation to an agent loss demands more than robustness to a single link loss, as the loss of one agent implies the loss of all control links incident to it. Therefore, we focus on the multiple agent loss problem throughout this paper. The issue can be dealt with in many different ways [15]. However, in this paper we incorporate a *proactive* approach. We introduce the robustness into the information architecture *a priori* , by using *redundant links*, in order to mitigate the effect of multiple agent loss. The measure of robustness here, is the number of agents the formation can afford to lose while preserving its cohesiveness.

We assume a graphical abstraction of the multi-agent formation, via an undirected graph $G = (V, E)$, in which each vertex corresponds to an agent and an edge corresponds to a bidirectional distance control law, actively maintaining the distance between the corresponding agents. This enables us to study the robustness of these formations via *rigidity* of the formation graph (an area of graph theory that deals with the characterization of graphs corresponding to formations which are rigid). Generally speaking, a formation, and by extension its underlying graph, is termed rigid if the distance between each pair of agents remains constant over time, normally through a subset of the inter-agent distances being actively maintained at prescribed values. It has been shown that this property is generic, in the sense that the rigidity of almost all formations with the same graph depends only on the topology of its graph and not the actual distances between the agents. This implies that having enough well-distributed control links within the formation will lead to a rigid formation no matter what the actual distances are between agents.

The rigidity of graphs has been extensively studied in the past [3], especially those corresponding to formations in an ambient two-dimensional space. However, robustness to multiple agent/link loss is a very new topic. In [17], robustness is defined by introducing the notion of k-edge/k-vertex rigidity of a graph: the graph remains rigid after the loss of up to $k-1$ edges/vertices; [17] also derived a collection of general properties of such graphs. In [18], a characterization is derived for k-edge rigidity. The author in [13] introduced

This work is supported by USAF-AOARD-10-4102. S. Alireza Motevallian is supported by National ICT Australia (NICTA). The work of C. Yu was supported by the Australian Research Council through a Queen Elizabeth II Fellowship, Discovery Project DP-110100538 and Overseas Expert Program of Shandong Province. Brian D.O. Anderson is supported by the ARC and NICTA.

S. Alireza Motevallian and Brian D.O. Anderson are with the Research School of Information Sciences and Engineering, the Australian National University and National ICT Australia.

C. Yu is with the Australian National University and NICTA Ltd, Canberra, A.C.T., Australia, and also with Shandong Computer Science Center, Jinan, China.

Emails: {alireza.motevallian,brad.yu, brian.anderson}@anu.edu.au.

Correspondence to: S. Alireza Motevallian, RSISE building 115, The Australian National University, Canberra, ACT, 0200, Australia. Tel: +61 2 6125 8639. Fax: +61 2 6125 8660.

978-1-61284-801-3/11/$26.00 ©2011 IEEE

birigid graphs, which are graphs remaining rigid after the loss of one vertex. In a different direction [15] proposes a procedure to restore the rigidity property of a rigid formation, when one of its agents is lost.

The main contribution of this paper is to parallel the path posed in [17], [18] and [15], focusing on the loss of more than one agent. In a related study [10], we have derived some distinct necessary and sufficient conditions for graphs which are k-vertex *globally rigid*, a stronger condition than rigidity. The similar results for rigidity will be provided here as a partial characterization of k-vertex rigid graphs, i.e. graph which remain rigid after the loss of up to $k-1$ vertices. As it turns out, for $k \leq 3$ a *size independence property* holds, i.e. the cost, in terms of *extra edges* required to secure the k-vertex rigidity property as opposed to mere rigidity, is very small and independent of the number of agents in the formation. As is later explained, this low cost property is lost when robustness against the loss of more than 2 agents is sought. This observation reinforces our focus on 3-vertex rigid graphs with the minimum edge count. A partial characterization of these graphs will be derived, followed by a constructive approach through which one can obtain a 3-vertex graph on arbitrary number of vertices. This characterization enables the formation designer to embed the robustness into a formation prior to any actual mission.

The structure of the paper is as follows. In Section II, some basic definitions are provided. Section III contains the main contribution of the paper and is followed by the concluding remarks in Section IV.

II. Background

Rigidity theory has been well studied in the literature [3], [16]. In this section, we introduce some basic definitions and properties required for the main contribution. The interested reader may refer to [3] for further details. Throughout this paper we assume that the formation lies in an ambient two-dimensional space and is modeled by a graph whose rigidity implies the rigidity of the formation and all generic formations with the same graph [3].

A. *Rigidity* and *Minimal Rigidity*

As mentioned earlier, a graph is called rigid if the only smooth motions are those corresponding to translation and rotation of the whole formation (see [3] for a precise definition). According to [16], a graph is minimally rigid if it is rigid and removing any one of the edges results in a nonrigid graph (for which any corresponding formation would have the possibility of flexing motions, apart from translation and rotation). There is a celebrated theorem, called *Laman's Theorem* [6], which gives a fully combinatorial characterization of minimally rigid graphs. In [6], it is also proved that every rigid graph contains a minimally rigid subgraph with the same vertices.

B. *Redundant Rigidity*

A graph is said to be *redundantly rigid*, if it is rigid and after removing any of the edges it still remains rigid. This notion can be generalized to the loss of k edges and/or k vertices. A graph will be termed *k-edge rigid* if after deletion of any set of up to $k-1$ edges, a rigid graph always results. With the same notion a graph is *k-vertex rigid* if after deletion of any set of up to $k-1$ vertices, the resulting graph is still rigid. As mentioned before, we are interested in addressing multiple agent loss. Therefore, in the remainder of the paper we study k-vertex rigidity of the underlying graphs. The following theorem and lemma were first proved in [17] and due to their relevance are restated here (we provide the proof of Lemma 2 as it summarizes several observations in [17]):

Theorem 1. *If $G = (V, E)$ is a k-vertex rigid graph, then each vertex has a degree of at least $k+1$.*

Lemma 2. *In a k-vertex rigid graph $G = (V, E)$, $|V| \geq k+3$ except for K_{k+2}, the complete graph on $k+2$ vertices.*

Proof: Suppose to obtain a contradiction that $|V| \leq k+2$ and G is not K_{k+2}. The k-vertex rigidity of G implies $\delta_G(v) \geq k+1$ (degree of vertex $v \in V$) which implies that $|V|$ cannot be less than $k+2$. However, the only graph with $|V| = k+2$ in which $\delta_G(v) \geq k+1$ holds is K_{k+2}. This contradiction implies $|V| \geq k+3$. ■

A graph is called *minimally* k-vertex (k-edge) rigid, if it is k-vertex (k-edge) rigid but after removing any one of the edges the resulting graph is no longer k-vertex (k-edge) rigid.

In the case of mere rigidity, the definition of minimal rigidity is proved to be equivalent to an alternative statement: a rigid graph is called minimally rigid if it has the *minimum number of possible edges* among all rigid graphs with the same number of vertices. Unfortunately, for general redundant rigidity (k-edge rigidity or k-vertex rigidity), these two notions are no longer equivalent; there are some graphs which are minimally redundantly rigid but the number of edges is not the minimum possible one among such graphs with the same vertex count. This property of redundant rigidity leads us to two different notions: strongly minimal and weakly minimal redundant rigidity [15].

- A k-vertex (k-edge) rigid graph is said to be *strongly minimal* if it has the minimum possible number of edges on a given number of vertices.
- A k-vertex (k-edge) rigid graph is said to be *weakly minimal* if it has more than the minimum possible number of edges on a given number of vertices, but has the property that removing any edge destroys k-vertex (k-edge) rigidity.

Later, we will need the characterizing conditions for 2-vertex rigidity as an special case. The remainder of this section contains the results to characterize strongly minimal 2-vertex rigid graphs [13]. Figure 1, originally from [13], depicts examples of *strongly minimal* 2-vertex rigid graphs.

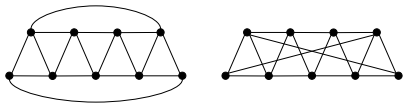


Figure 1. Examples of the 2 possible partitions of the edge set for strongly minimal 2-vertex rigid graphs: (a) the degree 3 vertices are adjacent, (b) the degree 3 vertices are non-adjacent

Lemma 3. *If $G = (V, E)$ is a 2-vertex rigid graph on 5 or more vertices, then $|E| \geq 2|V| - 1$.*

Theorem 4. *Let $G = (V, E)$ be a strongly minimal 2-vertex rigid graph on 5 or more vertices. Then G has exactly 3 vertices with degree 3 and the remaining vertices have degree 4, which implies $|E| = 2|V| - 1$.*

Theorem 5. *A graph $G = (V, E)$ is strongly minimal 2-vertex rigid if and only if G has exactly two vertices of degree 3 and there is a partition of the edge set E*

$$E = E_1 \cup E_2 \cup ... \cup E_k$$

such that the graph induced by $E \backslash E_i$ is minimally redundantly *rigid (i.e. the removal of any edge destroys redundant rigidity) for all i, and either*

- E_1 and E_2 are the edges incident to the two non-adjacent vertices of degree 3, respectively, and E_i is a single edge for $3 \leq i \leq k$, or
- E_1 is the union of the edges incident to the two adjacent vertices of degree 3, and E_i is a single edge for $2 \leq i \leq k$.

III. Results

In this section we start by studying a sufficient condition for k-vertex rigidity (Section III-A). This will enable us to propose formation structures which are robust against the loss of up to $k-1$ agents. Then, from a different perspective, the structure of k-vertex rigid graphs is studied and a distinct necessary condition is obtained (Section III-B). We will show the size independence property for $k \leq 3$ and display the optimality of structures which are 3-vertex rigid (highest value of k with size independence property). Therefore, we will some characterization of the special case of 3-vertex rigidity. It will be done by introducing some results on the relation between vertex and edge counts and vertex degrees of these graphs. Starting from $|V| = 6$, we will propose a constructive approach which can grow such graphs to arbitrary size via an extension operation.

A. Sufficient Condition for k-Vertex Rigidity

The notion of k-connectivity has been well studied in the literature and there are efficient algorithms to check this property on a given graph [11]. The idea here is to find a $(k+j)$-connectivity condition which is sufficient to k-vertex rigidity. The motivation comes from [8] where it is shown that in $2D$, 6-connectivity implies rigidity and 6 is the least possible number for this condition (k-connectivity with $k < 6$ is not sufficient for rigidity). The following theorem is derived from a stronger theorem we first proposed in [10].

We can extend this result to the case of k-vertex rigidity, as the following theorem shows:

Theorem 6. *Assume that $G = (V, E)$ is a $(k+5)$-connected graph. Then G is k-vertex rigid.*

The importance of this result is that it shows that the more complicated and awkward-to-verify property of k-vertex rigidity can be reduced to a well-known and easier to study property of k-connectivity.

B. Necessary Condition for k-Vertex Rigidity

In this subsection we derive a necessary condition similar to what we have derived in [10] but for k-vertex rigidity. This result gives a lower bound on the number of edges of such graphs. From [6], [13] we know that the minimum required number of edges for a graph to be rigid and 2-vertex rigid are $|E| = 2|V| - 3$ and $|E| = 2|V| - 1$, respectively. This observation leads us to an important question: since in both the above lower-bounds the number of edge is ***twice*** the number of vertices ($|V|$) plus a constant, can we conjecture that for any $k \geq 2$, the edge count of any strongly minimal *k-vertex rigid* graph satisfies the condition $|E| = 2|V| + c(k)$ (c is independent of $|V|$)? Unfortunately, the answer is no and, as we show in Theorem 8, this property is only valid for $k \leq 3$. For such values of k we say that the k-vertex rigid graph, has the ***size independence property***, meaning that the number of edges required to obtain a k-vertex rigid graph from a rigid graph is independent of the ***number of vertices*** ($|V|$).

Theorem 7 gives a lower bound on the edge count of a k-vertex rigid graph. We should mention that the same result is valid for k-edge rigid graphs as well.

Theorem 7. *In a strongly minimal k-vertex rigid graph, with $k \geq 2$, the edge count is under-bounded by the formula $|E| \geq \lceil \frac{k+1}{2} |V| \rceil + c(k)$, where $c(k)$ is an integer (c is independent of $|V|$ but depends on k) and for $k \geq 3$, if the equality holds (i.e.$|E| = \lceil \frac{k+2}{2} |V| \rceil + c(k)$), then $c(k) \geq 0$.*

Proof: Assume that $G = (V, E)$ is a strongly minimal k-vertex rigid graph of $|V| \geq k+3$ vertices, whose number of vertices is $|E| = a|V| + c(k)$ (according to Theorem 11 in [10] such a graph exists), where $c(k)$ is independent of $|V|$. According to Theorem 1, $\delta_i \geq k+1$ holds. Therefore, the average degree in G is $\delta_{avg} \geq k+1$. On the other hand, $\delta_{avg} = \frac{2|E|}{|V|} = 2a + \frac{2c(k)}{|V|}$. Hence, $2a + \frac{2c(k)}{|V|} \geq k+1 \Rightarrow k \leq (2a-1) + \frac{2c(k)}{|V|}$.

Since the property must hold for graphs of arbitrary size and in particular arbitrarily large $|V|$, assuming$|V| > 2(c(k)$, we will have $k \leq 2a-1$ or $a \geq \frac{k+1}{2}$. Therefore, $|E| \geq \lceil (\frac{k+1}{2}) |V| \rceil + c(k)$ holds for $|V| > 2c(k)$.

Now suppose that for some $|V|$ the equality holds (i.e. $|E| = \lceil \frac{k+1}{2} |V| \rceil + c(k)$). We prove that for such a strongly minimal k-vertex globally rigid graph $c(k) \geq 0$ always holds.

First suppose k is odd. Then, we will have $\delta_{avg} = k+1+\frac{2c(k)}{|V|} \geq k+1$, which implies $c(k) \geq 0$.

If k is even, then $|E| = \frac{k}{2}|V| + \lceil \frac{|V|}{2} \rceil + c(k)$ which gives $|E| < \frac{k}{2}|V| + c(k) + \frac{|V|}{2} + 1$. Therefore, $\delta_{avg} < k + \frac{2c(k)}{|V|} + 1 + \frac{2}{|V|}$ holds. This implies $k+1 < k+1+\frac{2c(k)}{|V|} + \frac{2}{|V|}$ which gives $-1 < c(k)$ and so $c(k) \geq 0$ holds. ∎

Theorem 8. *The highest value of k which the k-vertex rigid graph $G_k = (V, E)$ can have the size independence property is $k = 3$.*

Proof: As is shown in the proof of Theorem 7, the inequality $k \leq 2a-1$ holds. On the other hand, if G_k has the size independence property, then we necessarily must have $a = 2$ (as for rigidity and 2-vertex rigidity $a = 2$). This implies $k \leq 3$. Therefore $argmax_k(G_k) = 3$ holds for *size independent* graphs G_k and the proof is complete. ∎

From this result we can conclude that a formation designer can embed the robustness to the loss of up to ***two*** agents into a formation of any agent count, by adding only a fixed set of control links with possible redistribution of some edges to be incident on different vertex pairs. Actually by studying 3-vertex rigid graphs in the next section, we will derive the exact number of required control links to obtain such a robustness level (Lemma 9).

The size independence property is not true when there is robustness to the loss of more than ***two*** agents. Table I shows an example of a formation with 100 agents and the required number of control links for any desired level of robustness. As is evident, there is a considerable increase in the required number of control links when we want to obtain robustness to the loss of more than ***two*** agents.

# of agent losses tolerated	# of required links
0	197
1	199
2	202
3	$250 + c^+$

Table I

THE REQUIRED NUMBER OF CONTROL LINKS IN A 100-AGENT FORMATION WITH DIFFERENT ROBUSTNESS PROPERTIES. c^+ IS A POSITIVE INTEGER.

C. Relations between edge count, vertex count and vertex degrees

In this subsection, we study 3-vertex rigidity by presenting some results linking the vertex and edge counts and vertex degrees. The

first result underbounds the edge count in a 3-vertex rigid graph in terms of the vertex count.

Lemma 9. *If $G = (V, E)$ is a 3-vertex rigid graph with $|V| \geq 6$, then $|E| \geq 2|V| + 2$.*

Proof: First we prove that the inequality $|E| \geq 2|V| + 1$ holds. It is then enough to show that there is no 3-vertex rigid graphs on $|E| = 2|V| + 1$ edges.

For the inequality, to obtain a contradiction, suppose G is a 3-vertex rigid graph and has $|E| \leq 2|V|$ edges. In this case the average vertex degree is at most 4. Since all vertex degrees must be at least 4, this implies that the degree of each vertex is precisely 4, so that $|E| = 2|V|$.

Now consider $v_1, v_2 \in V$ of degree 4 which are not adjacent to each other (and always exist since $|V| \geq 6$) , and remove them from G to produce a graph $G' = (V', E')$. Then $|E'| = 2|V| - 8 = 2|V'| - 4$. Obviously G' is not rigid, which contradicts the fact that G is 3-vertex rigid.

To prove that $|E| = 2|V| + 1$ is impossible, suppose that the equality holds and set $n = |V|$. The average vertex degree in G is $\frac{\sum \delta_i}{n} = \frac{2|E|}{n} = \frac{4n+2}{n} = 4 + \frac{2}{n} > 4$. Therefore, there is at least one vertex with degree of more than 4. Now suppose $v_1, v_2 \in V$ with degrees k_1 and k_2, respectively. Observe that $G - v_1 - v_2$, obtained by removing v_1, v_2 and all their incident edges from G, has $n-2$ vertices and at most $|E| = 2n + 1 - (k_1 + k_2 - 1) = 2(n-2) - (k_1 + k_2 - 6)$ edges (the bound being achieved when v_1, v_2 are neighbors). Since $G - v_1 - v_2$ is rigid, $|E| \geq 2(n-2) - 3$ must hold. So $k_1 + k_2 - 6 \leq 3$ or $k_1 + k_2 \leq 9$. By considering $k_1 \geq 4$, $k_2 \geq 4$ we conclude that $4 \leq k_1 \leq 5$ and $4 \leq k_2 \leq 5$. Finally, if there are m vertices of degree 5, we have $5m + 4(n - m) = 2(2n + 1)$, which gives $m = 2$. This means that such a graph (if it exists) should have exactly two vertices of degree 5 which are adjacent and the others with degree 4.

Now we prove that such a graph cannot be 3-vertex rigid. First consider the case that $n = 6$. In this case all vertices of degree 4 are connected to vertices of degree 5 (figure 2). It is obvious that by removing vertices with degree 5, the resulting graph is not rigid. Now consider $n \geq 7$. In this case there are at least one pair (v_1, v_2) with degrees 4, 5 which are not adjacent. $G - v_1 - v_2$ has $2n + 1 - 5 - 4 = 2n - 8 = 2(n-2) - 4$ edges which contradicts the fact that it is rigid. This completes the proof. ∎

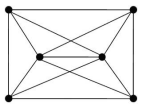


Figure 2. A graph with two 5-degree and four 4-degree vertices.

Figure 3 shows an example of a strongly minimal 3-vertex rigid graph with $|E| = 2|V| + 2$ on 6 vertices. Later in Section III-D, we will show that this class of graphs is an infinite set, by introducing an *extension operation* which can grow a strongly minimal 3-vertex rigid graph by one vertex. In the following theorem an in anticipation of the demonstration of existence of such graphs for arbitrary $|V| \geq 6$, we present a condition on the vertex degrees of these graphs.

Theorem 10. *There exists a strongly minimal 3-vertex rigid graph $G = (V, E)$ with $|E| = 2|V| + 2$ for any $|V| \geq 6$. In G, there are exactly 4 vertices with degree of 5 which are all adjacent (forming a K_4 subgraph) and all other vertices have degree of 4.*

Proof: This graph has an average degree of $4 + \frac{4}{n}$ (taking $|V| = n$). Therefore, there is at least one vertex with degree of more than 4. Now suppose $v_1, v_2 \in V$ with degrees k_1, k_2, respectively. Then, $G - v_1 - v_2$ has $n - 2$ vertices and at most $|E(G - v_1 - v_2)| = 2n + 2 - (k_1 + k_2 - 1) = 2(n-2) - (k_1 + k_2 - 7)$ edges (when v_1, v_2 are neighbors). Since $G - v_1 - v_2$ is rigid, $|E(G - v_1 - v_2)| \geq 2(n-2) - 3$ holds. Hence, $k_1 + k_2 - 7 \leq 3$ or $k_1 + k_2 \leq 10$. By considering $k_1 \geq 4$, $k_2 \geq 4$ we conclude that $4 \leq k_1 \leq 6$ and $4 \leq k_2 \leq 6$.

Finally, if there are m vertices of degree 6 and t vertices of degree 5, we have $6m + 5t + 4(n - m - t) = 2(2n + 2)$, which gives $2m + t = 4$. There are 3 possible cases:

a. $m = 1$, $t = 2$: in this case if we remove the 6-degree vertex in addition to a 5-degree one, the number of edges will be at most $|E(G - v_1 - v_2)| = 2n + 2 - (6 + 5 - 1) = 2n - 8 = 2(n - 2) - 4$ which contradicts the fact that $G - v_1 - v_2$ is rigid and $|E(G - v_1 - v_2)| \geq 2(n-2) - 3$. Therefore, this case cannot occur.

b. $m = 2$, $t = 0$: with the same argument as the case **a**, by removing two vertices with degree of 6 we have $|E(G - v_1 - v_2)| = 2n + 2 - (6 + 6 - 1) = 2n - 9 = 2(n-2) - 9$ and it is obvious that the resulting graph is not rigid. Hence again, this case cannot occur.

c. $m = 0$, $t = 4$: The proof of this case is trivial. The only important condition is that all 5-degree vertices should be adjacent. Otherwise, removing any 2 of them results in a non-rigid graph ($|E(G - v_1 - v_2| = 2(n-2) - 4 < 2(n-2) - 3$). Figure 3 shows such a graph with 6 vertices. ∎

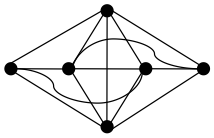


Figure 3. Strongly Minimal 3-vertex rigid graph with $|V| = 6$.

Remark 11. Observe that the condition provided by theorem 10 is not sufficient to ensure strongly minimal 3-vertex rigidity. As a counter example, consider the graph G shown in Figure 4 (left side). It is easy to observe that this graph satisfies Theorem 10. However, as shown in the right side of the figure, removing vertices shown by unfilled circles results in a non-rigid graph. Therefore, G is not 3-vertex rigid.

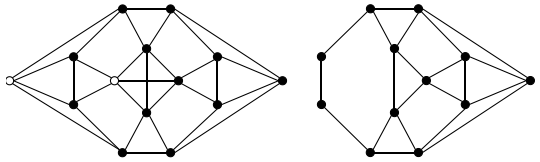


Figure 4. No 3-vertex rigid graph which satisfies Theorem 10

D. Growing strongly minimal 3-vertex rigid graphs

In this section we will show an infinite class of strongly minimal 3-vertex rigid graphs. The approach is to propose an extension operation which can be applied in any strongly minimal 3-vertex rigid graph and increases its vertex count by one. This operation is a special case of *X-replacement* operation [15] used before for extending weakly minimal 2-vertex rigid graphs - we call this new one *4-5 X-Replacement*, for reasons which are about to become apparent. Suppose that the original graph is $G = (V, E)$. Choose two edges $e_1 = \{a, b\}$ and $e_2 = \{c, d\}$ and $e_1, e_2 \in E$ so that a, c have degree 4 and are non-adjacent and b, d have degree 5 (that such a choice is in fact possible is proved below). Remove e_1, e_2 and add a new vertex called z. Connect z to a, b, c, d. In *4-5 X-Replacement* operation the degree of the original vertices remains the same and a new vertex of degree 4 is added

to G. Therefore, the new graph satisfies the conditions of the above theorems.

It remains to prove that this operation preserves strongly minimal 3-vertex rigidity. First we provide a lemma which guarantees that there are always appropriate edge choices for *4-5 X-replacement* operation.

Lemma 12. *If $G = (V, E)$ is a strongly minimal 3-vertex rigid graph, there are always two vertices, a, c say, with degree of 4 which are not adjacent, yet are connected to two different vertices, b and d say, with degree of 5.*

Proof: Since the vertices of degree 4 which are connected to the core (vertices of degree 5) can only be adjacent to at most 3 other vertices of the same degree 5, if the number of vertices of degree 4 is more than 4, there are at least two of them which are not adjacent. Therefore, the theorem is proved for $|V| > 8$. Since $|V| \geq 6$ holds in general, we need to prove the lemma for $|V| = 6, 7$ and 8. From Figure 3 it is evident that in the only strongly minimal 3-vertex rigid graph with 6 vertices, the only two 4-degree vertices are not adjacent. Finally, it is trivial to see that (up to isomorphism) the graphs in Figure 5 are the only possible strongly minimal 3-vertex rigid graphs of size 7 and 8 respectively. ∎

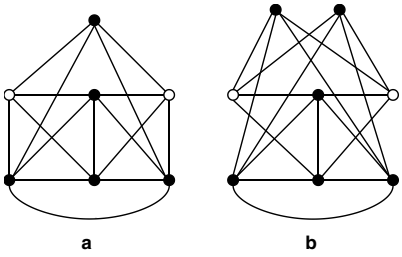


Figure 5. Strongly minimal 3-vertex rigid graphs of size 7 (a) and 8 (b). Vertices depicted by unfilled circles are 4-degree but not adjacent.

Theorem 13. *Suppose $G = (V, E)$ is a strongly minimal 3-vertex rigid graph. After applying the* 4-5 X-Replacement *operation on G, the resulting graph G' is strongly minimal 3-vertex rigid on $|V| + 1$ vertices.*

Proof: In this proof we show that by removing any pair of vertices from the graph G' after *4-5 X-replacement*, the resulting graph is still rigid. Suppose (a, b) and (c, d) are two candidate edges to be removed from G in *4-5 X-replacement*, with $\delta_a = \delta_c = 4$ and $\delta_b = \delta_d = 5$. The new vertex to be connected to all of a, b, c, d is e (Figure 6).

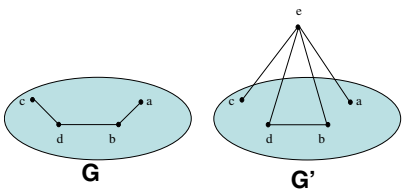


Figure 6. *4-5 X-replacement* operation. b, d have degree 5 while a, c have degree 4.

There are 6 distinct cases based on different choices for the vertex pair to be removed from G'. For further reference we call these vertices x and y.

a. $x, y \in V \backslash \{a, b, c, d\}$: We know that $G - x - y$ is rigid. In this case, none of the edges connecting e to one of $\{a, b, c, d\}$ is removed. Therefore, one can interpret $G' - x - y$ as an extension of $G - x - y$ by an ordinary *X-replacement* operation. This has been proved to preserve rigidity of $G' - x - y$ [15]. Hence, in this case the result of the operation is still a rigid graph.

b. $x \in \{a, b\}$ (or $x \in \{c, d\}$ which are the same by the symmetry property) and $y \in V \backslash \{a, b, c, d\}$: Without loss of generality suppose that for this case $x = a$ (Figure 7). Therefore, $G' - x - y$ can be obtained from $G - x - y$ by a standard edge splitting operation which preserves rigidity [16]. Since $G - x - y$ is rigid, $G' - x - y$ is also rigid.

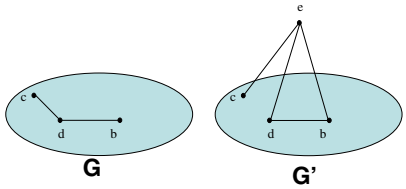


Figure 7. Removing $x = a$ and $y \in V \backslash \{a, b, c, d\}$ from G and G'. The operation reduces to the edge splitting operation on (c, d) edge.

c. $x \in \{a, b\}$ and $y \in \{c, d\}$ (one vertex from each edge). Again suppose that -without loss of generality- $x = a$ and $y = c$ (Figure 8). One can simply observe that $G' - x - y$ can be produced by applying a vertex addition to $G' - x - y$ which preserves rigidity when applied to a rigid graph [16]. Therefore, $G' - x - y$ is rigid.

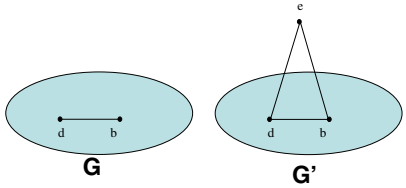


Figure 8. Removing $x = a$ and $y = c$. The operation reduces to vertex addition operation in which e is connected to b, d by 2 edges.

d. $x = a$ and $y = b$ (both are the extremes of one edge): Without loss of generality suppose that $x = a$ and $y = b$ (Figure 9). After removing vertex b, the resulting graph ($G' = G - b$) is still 2-vertex rigid and has $|E'| = 2|V| + 2 - 5 = 2|V'| - 1$ edges. According to theorem 4 and 5, it is strongly minimal 2-vertex rigid with exactly two vertices of degree 3 (which had been 4-degree vertices connected to b before it was removed). Since vertex a is one of the 3-degree vertices in this graph, according to theorem 5, removing it and all of its incident edges ensures that the resulting graph is *redundantly rigid*. Therefore, $G - x - y - (c, d)$ is still rigid. One can easily conclude that adding e with two edges to this graph, which forms $G' - x - y$, is a case of vertex addition extension of a rigid graph and therefore $G' - x - y$ is rigid.

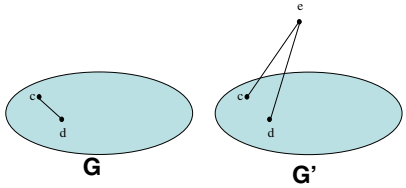


Figure 9. Removing $x = a$ and $y = b$.

e. $x = e$. In this case if $y \in \{a, b, c, d\}$, we can observe that the difference between $G - y$ and $G' - x - y$ is that $G - y$ has an extra

edge (c,d). Since $G-y$ is 2-vertex rigid, we can conclude that it is 2-edge rigid [18]. Therefore, by removing one edge like (c,d) from it, the result $(G'-x-y)$ is still rigid (Figure 10).

If $y \notin (a,b,c,d\}$, the difference between $G'-x-y$ and $G-y$ is in two edges (c,d) and (a,b) (Figure 11). From theorem 5 we know that $G-y-(c,d)$ is minimally redundantly rigid (since $E_i=(c,d)$ in one possible partition of $G-y$). Therefore, $G-y-(c,d)-(a,b)$ is rigid. By looking at Figure 11 one can easily realize that $G-y-(c,d)-(a,b)$ is $G'-y-x$. Hence, $G'-y-x$ is rigid.

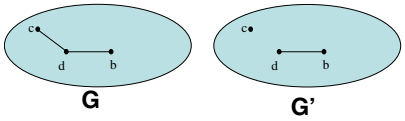


Figure 10. Removing $x=e$ and $y=a$.

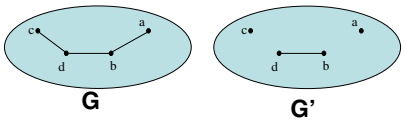


Figure 11. Removing $x=e$ and $y \in G\backslash\{a,b,c,d\}$

This completes the proof of Theorem 13. ■

IV. Conclusions

In this paper we studied the robustness of information architectures to control the formation of autonomous agents. As these agents sometimes operate in hazardous environments like battle-fields, they are prone to multiple agent and/or link loss. Due to the higher severity of agent loss than link loss, the main contribution of this paper is to propose information architectures for multi-agent formations, which are robust against the loss of multiple agents, in the sense of retaining rigidity of the formation when such loss occurs. By adopting the notion of k-vertex rigid graphs (corresponding to formations which are tolerant to the loss of up to $k-1$ agents, we derived some distinct necessary and sufficient conditions. We also established the *size independence property* for $k \leq 3$, i.e. the cost, in terms of *extra edges* required to secure the k-vertex rigidity property as opposed to mere rigidity, is very small and independent of the number of agents in the formation. As explained before, this low cost property is lost when robustness against the loss of more than two agents is sought. This observation motivated the detailed study of 3-vertex rigid graphs with the minimum edge count. A partial characterization of these graphs was derived, followed by a constructive approach through which one can obtain a 3-vertex rigid graph on arbitrary number of vertices. This characterization enables the formation to be designed with the robustness against the loss of up to two agents prior to any actual mission, a proactive approach.

The main assumption in this paper was that the control operation is cooperative in the sense that both agents at the ends of a control link, cooperate to maintain the desired distance. This enabled us to model the formation with an undirected graph. However, there are other schemes in which only one of the agents is responsible to maintain a distance. These problems and the robustness of such formations can be studied via the definition of *persistent* formations [2] as an extension of this work.

Finally, although we provided some necessary and sufficient conditions on k-vertex rigid formations, it is still an open problem to fully characterize such architectures. This is one possibility for future work.

References

[1] B. Anderson, B. Fidan, C. Yu, and D. Walle. UAV formation control: Theory and application. *Recent Advances in Learning and Control*, pages 15–33, 2008.

[2] B. Anderson, C. Yu, B. Fidan, and J. Hendrickx. Rigid graph control architectures for autonomous formations. *Control Systems Magazine, IEEE*, 28(6):48–63, 2008.

[3] J.E. Graver, B. Servatius, and H. Servatius. *Combinatorial rigidity*. American Mathematical Society, 1993.

[4] Daniel L. Haulman. U.s. unmanned aerial vehicles in combat. Technical report, Air Force Historical Research Agency, 2003.

[5] T. Kopfstedt, M. Mukai, M. Fujita, and C. Ament. Control of Formations of UAVs for Surveillance and Reconnaissance Missions. In *17th IFAC World Congress, COEX, South Korea*, pages 6–11, 2008.

[6] G. Laman. On graphs and rigidity of plane skeletal structures. *Journal of Engineering Mathematics*, 4:331–340, 1970.

[7] Jay Levine. Nasa j-ucas x-45a represents an evolutionary step in technology and a look at future aviation. *The X-press, Dryden Flight Research Center*, 46(7), 2004.

[8] L. Lovasz and Y. Yemini. On generic rigidity in the plane. *SIAM Journal on Algebraic and Discrete Methods*, 3:91, 1982.

[9] R.K. Mehra, J.D. Boskovic, and S.M. Li. Autonomous formation flying of multiple UCAVs under communication failure. In *Position Location and Navigation Symposium, IEEE 2000*, pages 371–378. IEEE, 2000.

[10] S. Alireza Motevallian, Changbin Yu, and Brian D. O. Anderson. Robustness to the loss of multiple nodes in the localizability of sensor networks. To be published in IFAC WC 2011, August 2011.

[11] H. Nagamochi and T. Ibaraki. *Algorithmic aspects of graph connectivity*. Cambridge University Press New York, NY, USA, 2008.

[12] T. Nisser and C. Westin. Human factors challenges in unmanned aerial vehicles (uavs): A literature review. *Lund University School of Aviation, Tech. Rep. TFHS*, 5:1, 2006.

[13] B. Servatius. Birigidity in the plane. *SIAM Journal on Discrete Mathematics*, 2(4):582–599, 1989.

[14] R.S. Smith and F.Y. Hadaegh. Control of deep-space formation-ïňĆying spacecraft: relative sensing and switched information. *AIAA Journal of Guidance, Control, and Dynamics*, 28(1):106–114, 2005.

[15] T.H. Summers, C. Yu, and B. Anderson. Addressing agent loss in vehicle formations and sensor networks. *International Journal of Robust and Nonlinear Control*, 19(15):1673–1696, 2009.

[16] T.S. Tay and W. Whiteley. Generating isostatic frameworks. *Structural Topology 1985 núm 11*, 1985.

[17] C. Yu and B. D. O. Anderson. Development of redundant rigidity theory for formation control. *International Journal of Robust and Nonlinear Control*, 19(13):1427–1446, 2009.

[18] C. Yu, S. Dusgupta, and B. D. O. Anderson. Network localizability with link or node losses. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 402–407, 2010.

Preprints of the 18th IFAC World Congress
Milano (Italy) August 28 - September 2, 2011

Robustness to the Loss of Multiple Nodes in the Localizability of Sensor Networks ⋆

S. Alireza Motevallian * **Changbin Yu** **
Brian D.O. Anderson *

* *Research School of Information Sciences and Engineering, The Australian National University and National ICT of Australia (e-mail: {alireza.motevallian,brian.anderson}@anu.edu.au).*
** *Research School of Information Sciences and Engineering, The Australian National University (e-mail: brad.yu@anu.edu.au).*

Abstract: In the studies on the localization of wireless sensor networks (WSN), it has been shown that a network is in principle uniquely localizable if its underlying graph is globally rigid and there are at least $d+1$ non-collinear anchors (in d-space). The high possibility of the loss of nodes or links in a typical WSN, specially mobile WSNs where the localization often needs to be repeated, enforces to not only have localizable network structures but also structures which remain localizable after the loss of multiple nodes/links. The problem of characterizing robustness against the loss of ***multiple*** nodes, which is more challenging than the problem of multiple link loss, is being studied here for the first time, though there have been some results on ***single*** node loss. We provide some sufficient properties for a network to be robustly localizable. This enables us to answer the problem of how to make a given network robustly localizable. We also derive a lower bound on the number of the links such a network should have. Elaborating it to the case of robustness against the loss of up to 2 nodes, we propose the optimal network structure, in terms of the required number of distance measurements.

Keywords: Robust Localizability, Localization, Wireless Sensor Networks, Global Rigidity.

1. INTRODUCTION

Knowing the location of the nodes in a wireless sensor network is critical, as in many applications, the interpretation of the data and decision making is impossible without knowing the position of the detected event. The mobility or unplanned deployment of the nodes in a WSN necessitates a localization technique which can be frequently executed. A variety of techniques are proposed in the literature. Among them, there are schemes in which only a small number of special nodes (called *anchors*) have their positions known a priori (Aspnes et al., 2006), sometimes because they are GPS-equipped. Then by obtaining a set of distance measurements between enough pairs of ordinary nodes, an algorithm determines the node positions, using the distances and anchor location data; the process is called *network localization*.

There is a fundamental question in network localization that need to be answered prior to the localization process: what structure should a network have, in order to be localizable? It is important to see beforehand if the network is localizable, as it is a waste of effort to seek to localize a network which is actually not localizable. This question is answered in (Aspnes et al., 2006) with the help of recent results from graph theory including the concept of Global Rigidity. It is proved there that a network is uniquely localizable if and only if its underlying graph is globally rigid and there are at least $d+1$ non-collinear anchors in d-space ($d \in \{2,3\}$). The *underlying graph* $G(V,E)$ of a network is the one in which there is a vertex corresponding to each network node in the vertex set V and two vertices are connected via an edge in the edge set E if the distance between the corresponding nodes is known (Note that in modeling a network, the graph itself does not contain the length data). A ***realization*** of a graph with associated length data is an assignment of the vertices of the graphs to points in $\mathbb{R}^d$ such that the distance between points corresponding to adjacent vertices in the graph equals the distance associated with the corresponding edge of the graph. A graph is called globally rigid if all of its realizations in the d-space are congruent, i.e. can be obtained from another realization of the graph only by a combination of reflections, rotations and translations of the whole graph. It is a nontrivial result of the theory that global rigidity is a property determinable from the graph alone (i.e. without the distance set) provided that the distances correspond to generic node positions (e.g. collinearities are likely to be excluded).

One of the most challenging issues in sensor networks is the high possibility of failures either in communication links or sensor nodes themselves, due to different causes, e.g. signal jamming, obstacle, power depletion, mechanical failure, etc. This may result in changes to the structure

⋆ This work is supported by USAF-AOARD-10-4102. S. A. Motevallian is supported by National ICT Australia (NICTA). C. Yu is supported by the Australian Research Council (ARC) through a Queen Elizabeth II Fellowship and Discovery project DP-110100538. B.D.O. Anderson is supported by the ARC and NICTA.

of the network (and also in the underlying graph). In other words, such failures may cause a previously uniquely localizable network to become non-localizable; in a mobile network where re-localization must occur due to the motion, this is especially serious. It is obvious that coping with a single node loss is more demanding than a single link loss as removal of any node also results in the removal of all of its incident links. The solution to the problem of securing tolerance against node loss is to introduce some sort of redundancy in the distance measurements (links), i.e. having network structures which are robust against the loss of some nodes and/or communication links. Despite the importance of this robustness property, only those structures which are robust against the loss of a ***single*** node appear to have been studied up to now (Summers et al., 2008). In (Yu and Anderson, 2008) some general properties of *rigid* (a simpler property than *global rigidity*) which remain rigid after the loss of nodes/links are studied. This work is extended further in (Yu et al., 2010) for sensor network localization, concerning both node loss and link loss tolerance. The authors generalized the notion of redundant rigidity to (p, q)-rigidity: the ability to retain (global) rigidity given the loss of any $p-1$ nodes (including their incident links) and also any further $q-1$ links. However, they only established properties associated with redundancy (robustness) under the loss of any $q-1$ links ($(1, q)$-rigidity). As is argued in (Yu et al., 2010), characterizing $(p, 1)$-rigidity (redundancy under loss of any $p-1$ vertices) and (p, q)-rigidity are still open problems.

In this paper, the problem of robust localizability against the loss of ***multiple nodes*** in 2D is being studied for the first time, through proposing structures for underlying graph which remain globally rigid after the loss of p vertices. We also argue briefly in Section 3.1 that for networks defined by random geometric graphs (a common assumption for large scale sensor networks), formulae are available indicating a minimal transmission radius ensuring the robust localizability property. To elaborate on how the obtained results can be used, we study the case where $p = 2$ and suggest a class of graphs which are robust under the loss of up to 2 vertices. Such graphs are, as we show in the paper, optimal in terms of the number of distance measurements they requires. This is a starting point for further studies of such structures for general p.

In this work, we always assume the ambient space to be 2D, unless explicitly noticed. We also assume that there are $p+3$ anchors in the network so that after removing any p nodes there are still 3 anchors in the network. The structure of the paper is as follows. In Section 2, the required background is reviewed. Section 3 contains the main contribution of the paper. Finally, concluding remarks are made in Section 4.

2. BACKGROUND

In this section we recall some definitions and properties of (minimal) rigidity, redundant rigidity and global rigidity and followed by the general notion of redundant rigidity. For a formal definition and detailed introduction to rigidity and global rigidity please refer to (Graver et al., 1993; Connelly, 2005).

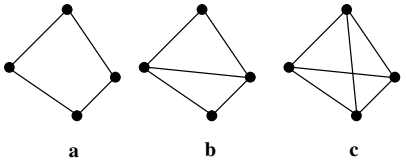


Fig. 1. (originally from (Yu and Anderson, 2009)) Realization of a graph in 2D that is (a) non-rigid, (b) minimally rigid and (c) redundantly rigid (also globally rigid).

2.1 Rigidity, Minimal Rigidity and Redundant Rigidity

Assume in a realization of an underlying graph in an ambient space (2D or 3D), each point is a revolute joint and each edge is a solid bar with an specified length. This *framework* (a term commonly used in rigidity theory), and therefore the underlying graph, is called rigid, if under any motion of the framework in the space, the distance between each pair of points remains constant over time, no matter whether there is or is not an explicit edge connecting them (see (Graver et al., 1993) for a more precise definition). The extension of the term *rigid* to refer to the graph is valid since it can be shown that if a realization of a graph is rigid for one set of length, a realization for almost any length set will also be rigid. A graph is minimally rigid if it is rigid and removing any one of the edges results in a nonrigid graph (Figure 1b). In (Laman, 1970), it is proved that every rigid graph contains a minimally rigid subgraph with the same vertex set. A graph is termed *redundantly rigid*, if it is rigid and after removal of any edge, it still remains rigid (Figure 1c).

2.2 Global Rigidity

A graph is termed globally rigid, if any two of its realizations in the space are congruent , i.e. each realization can be obtained from any other realization only by a combination of reflections, translations and/or rotations of the whole graph (Figure 1c). The following theorem from (Jackson and Jordan, 2005) gives a characterization of globally rigid graphs in 2D.

Theorem 1. The graph $G = (V, E)$ with $|V| \geq 4$ is globally rigid, if and only if it is 3-connected and redundantly rigid.

A graph is said to be *k-connected* if it is connected and after removal of *any* set of up to $k-1$ vertices, it still remains connected, see (Nagamochi and Ibaraki, 2008).

2.3 Generalizing Redundant (Global) Rigidity

The notion of redundant rigidity can be generalized to the loss of k edges and/or k vertices (Yu and Anderson, 2009). A graph is termed *k-edge rigid* if after deletion of any set of up to $k-1$ edges, a rigid graph always results. With the same notion a graph is *k-vertex rigid* if after deletion of any set of up to $k-1$ vertices, the resulting graph is still rigid. Similarly, a graph is called *k-edge* (*k-vertex) globally rigid* if after deletion of any set of up to $k-1$ edges (vertices), a globally rigid graph always results (Yu et al., 2010). The focus of this work is on k-vertex global

rigidity. In this paper, to avoid near trivialities and because the case is actually special in terms of the probable results, we assume that $G = (V, E)$ is **not** a complete graph (there is an edge between every vertex pair) as it is obvious that the complete graph K_l is $(l-2)$-vertex globally rigid and measuring the distances between all pairs of the vertices in the graph (which resulted in a complete underlying graph) is inefficient in practice.

Let us record a few properties of k-vertex globally rigid graphs. Lemma 2 is similar to Corollary 2 in (Yu and Anderson, 2009), but considers k-vertex global rigidity instead of k-vertex rigidity.

Lemma 2. If $G = (V, E)$ is k-vertex globally rigid, then it is $(k+2)$-connected. Also, each vertex $v \in V$ has degree at least $k+2$ ($\delta_G(v) \geq k+2$, where $\delta_G(v)$ is the degree of vertex v) and the number of vertices satisfies the inequality $|V| \geq k+3$.

Proof. To prove $(k+2)$-connectivity by obtaining a contradiction, suppose G is not $(k+2)$-connected. Then there exists a cut set, say S, where $|S| \leq k+1$. It is obvious (from the definition) that if $|V| < k-1$ then G cannot be k-vertex globally rigid. Now let $U \subset S$ be any arbitrary set of vertices with the condition that $|U| = k-1$. According to the definition of k-vertex global rigidity, $G \backslash U$ is still globally rigid. However, $T = S \backslash U$ is a cut-set in $G \backslash U$ and has $T = |S \backslash U| \leq 2$ vertices. This means that $G \backslash U$ is not 3-connected and therefore $G \backslash U$ is not globally rigid. This contradiction implies that G is $(k+2)$-connected.

The $(k+2)$-connectivity implies $\delta_G(v) \geq k+2$ for any $v \in V$. Otherwise, there is a set U of up to $k+1$ vertices (the neighbors of a vertex v with $\delta_G(v) \leq k+1$) whose removal makes the graph disconnected. It follows that$|V| \geq k+3$ holds as there should be at least $k+3$ vertices in V so that every vertex has a degree of at least $k+2$. ■

A graph is called *minimally* k-vertex globally rigid, if it is k-vertex globally rigid but after removing any one of the edges the resulting graph is no longer k-vertex globally rigid. It is not hard to argue that (a) k-vertex globally rigid graphs exists for any positive k (we present a particular construction in section 3.2 below) and (b) given any k-vertex globally rigid graph, a subgraph can be obtained by edge removal which is minimally k-vertex globally rigid. Thus the concept is not an empty one.

In (Jordan and Szabadka, 2009) it is shown that there is an operation, called *1-extension* (or *edge-splitting*), which can grow any globally rigid graph by 1 vertex in 2D. In this operation, an edge $(u, v) \in E$ is removed from E and a new vertex z, is connected to both u and v and an arbitrary third vertex $w \in V$, $w \notin \{u, v\}$.

Based on Theorem 1.2 in (Jordan and Szabadka, 2009), the number of edges in a minimally globally rigid graph (the graph $G = (V, E)$ which is globally rigid and $\forall e \in E$ $G' = (V, E - e)$ is not globally rigid) is:

Lemma 3. If $G = (V, E)$ is a minimally globally rigid graph on $|V| \geq 4$ vertices, then $|E| = 2|V| - 2$ always holds.

In the case of standard global rigidity, the definition of minimal global rigidity is proved (Theorem 1.2 (Jordan and Szabadka, 2009)) to be equivalent to an alternative statement: a rigid graph is called minimally globally rigid if it has the *minimum number of possible edges* ($|E| = 2|V| - 2$) among all globally rigid graphs with the same number of vertices. For k-vertex global rigidity, these two notions are no longer equivalent; there are some graphs which are minimally 2-vertex globally rigid but the number of edges is not the minimum possible among such graphs with the same vertex count. This property leads us to two different notions: *strongly minimal* and *weakly minimal* k-vertex global rigidity (this notion is adapted from the notion of strongly/weakly minimal 2-vertex *rigidity* in (Summers et al., 2008)).

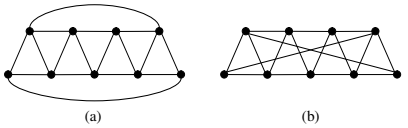


Fig. 2. (originally from (Servatius, 1989)) Examples of the 2 possible partition of edge set for strongly minimal 2-vertex rigid graphs: (a) the degree 3 vertices are adjacent, (b) the degree 3 vertices are non-adjacent

- A k-vertex globally rigid graph is said to be *strongly minimal* if it has the minimum possible number of edges (over all k-vertex minimally globally rigid graphs) on a given number of vertices.
- A k-vertex globally rigid graph is said to be *weakly minimal* if it has more than the minimum possible number of edges on a given number of vertices, but has the property that removing any edge destroys k-vertex global rigidity.

The following results of (Servatius, 1989) characterized the structure of strongly minimal 2-vertex rigid graphs for the first time. Due to their relevance and later use, we restate them here. Figure 2, illustrates examples of *strongly minimal* 2-vertex rigid graphs with each of the two possible types.

Lemma 4. (Lemma 1 of (Summers et al., 2008)) If $G = (V, E)$ is a 2-vertex rigid graph on 5 or more vertices, then $|E| \geq 2|V| - 1$.

Theorem 5. (Proposition 1 of (Servatius, 1989)) Let $G = (V, E)$ be a strongly minimal 2-vertex rigid graph on 5 or more vertices. Then G has exactly 3 vertices with degree 3 and the remaining vertices have degree 4, which implies $|E| = 2|V| - 1$.

Theorem 6. (Theorem 3.1 of (Servatius, 1989)) A graph $G = (V, E)$ with $|V| \geq 5$ is strongly minimal 2-vertex rigid if and only if G has exactly two vertices of degree 3 and there is a partition of the edge set E

$$E = E_1 \cup E_2 \cup ... \cup E_k$$

such that the graph induced by $E \backslash E_i$ is *minimally redundantly* rigid (i.e. the removal of any edge destroys redundant rigidity) for all i, and either

- E_1 and E_2 are the edges incident to the two non-adjacent vertices of degree 3, respectively, and E_i is a single edge for $3 \leq i \leq k$, or
- E_1 is the union of the edges incident to the two adjacent vertices of degree 3, and E_i is a single edge for $2 \leq i \leq k$.

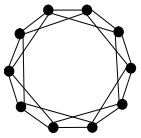


Fig. 3. A Strongly Minimal 2-Vertex Globally Rigid graph of size 10 (C_{10}^2).

The following theorem from (Summers et al., 2008) characterizes strongly minimal 2-vertex *globally* rigid graphs (Figure 3).

Theorem 7. (Theorem 10 in (Summers et al., 2008)) The graph $G = (V, E)$ of 5 or more vertices is strongly minimal 2-vertex globally rigid if and only if the following conditions hold:

- $|E| = 2|V|$
- G is 4-connected
- G is redundantly strongly minimal 2-vertex rigid

A redundantly strongly minimal 2-vertex rigid graph is one in which after removal of any edge, the result is strongly minimal 2-vertex rigid. Such a graph is obtained and can only be attained by joining the vertices of degree 3 in a strongly minimal 2-vertex rigid graph whose vertices of degree 3 are not adjacent. Figure 2 depicts examples of 2 possible configurations for strongly minimal 2-vertex rigid graphs. If in Figure 2(b), we connect the vertices of degree 3 (which are not adjacent), the result is a redundantly strongly minimal 2-vertex rigid graph.

3. RESULTS

In this section we start by studying some sufficient conditions for k-vertex global rigidity (Section 3.1). This will enable us to propose localizable structures which are robust against the loss of up to $k-1$ vertices. Then, from a different perspective, the structure of k-vertex globally rigid graphs is studied and some necessary conditions are obtained (Section 3.2). To elaborate this result, the case of 3-vertex global rigidity is considered in more detail and a general class of structures which are strongly minimal 3-vertex globally rigid are introduced (Section 3.3). We conclude this section by comparing these results on strongly minimal 3-vertex global rigidity and the result that can be obtained by the sufficient condition in Section 3.1.

3.1 Sufficient Condition for k-Vertex Global Rigidity

The notion of k-connectivity has been well studied in the literature and there are efficient algorithms to check this property on a given graph (Nagamochi and Ibaraki, 2008). The idea here is to identify a $(k+j)$-connectivity condition which is sufficient to ensures k-vertex global rigidity. It is proved in (Lovasz and Yemini, 1982) that in 2D, 6-connectivity implies rigidity and 6 is the least possible number for this condition (k-connectivity with $k < 6$ is not sufficient for rigidity). Recent extension to this work (Jackson and Jordán, 2009) showed the sufficiency of this condition for global rigidity as well.

Theorem 8. (restatement of Theorem 1.2 in (Jackson and Jordán, 2009)) Suppose $G = (V, E)$ is a 6-connected graph in 2D. Then it is 2-edge globally rigid (removing any edge results in a globally rigid graph).

We can extend this result to the case of k-vertex global rigidity, as the following theorem shows:

Theorem 9. Assume that $G = (V, E)$ is a $(k+5)$-connected graph. Then G is k-vertex globally rigid.

Proof. Let $U \subset V$ be any set of vertices with $|U| = k-1$. Since G is $(k+5)$-connected, it is obvious that $G' = G \backslash U$ is 6-connected ($k+5-(k-1) = 6$). Therefore, G' is globally rigid. Since U was chosen arbitrarily, we conclude that G is k-vertex globally rigid. ■

This result is very important as it reduces the more complicated and unfamiliar property of k-vertex global rigidity to a well-known and easier to study property of $(k+5)$-connectivity. Specifically for large-scale random wireless networks (where the nodes are assumed to be Poisson distributed with a certain density, and they have common transmission range), it is shown in (Wan and Yi, 2004) that asymptotically, i.e. as the node count tends to infinity, there is a relation between the transmission range and a k-connectivity property (for any k) of a wireless sensor network, i.e. by increasing the transmission power of the nodes in a random network above an specific threshold, the network becomes k-connected (The threshold is NOT of the same order as the diameter of the network, a situation which trivially would ensure k-connectivity). This idea tied to Theorem 9 answers the question of "*given a wireless sensor network, how can we make it robustly localizable?*":

Theorem 10. Assume a network of wireless sensor nodes. As the node count tends to infinity, there exists a critical transmission radius (NOT of the order of the network diameter), say r, such that by increasing the transmission range of every node above it, the network becomes k-vertex globally rigid (robustly localizable against the loss of up to $k-1$ vertices).

Finding the value of r for various k is beyond the scope of this paper.

3.2 Necessary Condition for k-Vertex Global Rigidity

In this subsection a necessary condition on k-vertex globally rigid graphs is provided which is a lower bound on the number of edges these graphs have. A prerequisite of the main theorem is to show that for fixed k, there always exists a k-vertex globally rigid graph of arbitrary size in which the number of edges depends linearly on the number of vertices.

Theorem 11. For ambient space dimension of $d = 2$ and 3, there exists a k-vertex globally rigid graph $G = (V, E)$ with $|V| \geq k+3$ and otherwise arbitrary, for which $|E| = a|V| + b$ holds, for some a and b dependent on k and d but independent of $|V|$.

Proof. The proof is by constructing a k-vertex globally rigid graph which satisfies the conditions of the theorem. First, observe that a complete graph K_{k+d+1} is k-vertex globally rigid. The number of edges in K_{k+d+1} is $m = |V(K_{k+d+1})| = \frac{(k+d)(k+d+1)}{2}$. Suppose $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are two K_{k+d+1} graphs. It is easy to show

that by adding a set L of $k+d+1$ edges between G_1 and G_2, such that every vertex in G_1 is connected to exactly 1 vertex in G_2 and conversely, $G = (V_1 \cup V_2, E_1 \cup E_2 \cup L)$ is also k-vertex globally rigid.

More generally, to construct a k-vertex globally rigid graph $G = (V, E)$ on $n = |V|$ vertices imagine a series of K_{k+d+1} graphs $\{G_i = (V_i, E_i)\}_{i=1..(p-1)}$ in which G_i is connected to G_{i+1} by $k+d+1$ edges, say L_i. Here, $p = \lfloor \frac{n}{m} \rfloor + 1$ and $|V(G_p)| = n - pm \leq m$. For G_p, which may have fewer than $k+d+1$ vertices, connect every vertex to all vertices of G_{p-1} (by $k+d+1$ edges), which constitute the set L_{p-1} of edges between G_{p-1} and G_p. It is easy to show that $G = (V, E) = (V_1 \cup ... \cup V_p, E_1 \cup ... \cup E_p \cup L_1 \cup ... \cup L_{p-1})$ is k-vertex globally rigid. Hence, the inequality $|E| \leq (p-1)m + (p-1)(k+d+1) + (k+d+1)m$ holds. With k fixed, $O(|E|) = O(pm) = O(n)$ holds, i.e. the number of edges is linear with respect to the number of vertices. ■

Theorem 12. In a strongly minimal k-vertex globally rigid graph the edge count is under-bounded by the formula $|E| \geq \lceil \frac{k+2}{2} |V| \rceil + c(k)$, where $c(k)$ is an integer (c is independent of $|V|$ but depends on k) and if the equality holds (i.e.$|E| = \lceil \frac{k+2}{2} |V| \rceil + c(k)$), then $c(k) \geq 0$.

Proof. Assume that $G = (V, E)$ is a strongly minimal k-vertex globally rigid graph of $|V| \geq k+3$ vertices, whose number of vertices is $|E| = a|V| + c(k)$ (according to Theorem 11 such a graph exists), where $c(k)$ is independent of $|V|$. According to Lemma 2, $\delta_i \geq k+2$ holds. Therefore, the average degree in G is $\delta_{avg} \geq k+2$. On the other hand, $\delta_{avg} = \frac{2|E|}{|V|} = 2a + \frac{2c(k)}{|V|}$. Hence, $2a + \frac{2c(k)}{|V|} \geq k + 2 \Rightarrow k \leq 2(a-1) + \frac{2c(k)}{|V|}$.

Since the property must hold for graphs of arbitrary size and in particular arbitrarily large $|V|$, assuming $|V| > 2(c(k)$, we will have $k \leq 2(a-1)$ or $a \geq \frac{k}{2} + 1$. Therefore, $|E| \geq \lceil (\frac{k}{2} + 1) |V| \rceil + c(k)$ holds for $|V| > 2c(k)$.

Now suppose that for some $|V|$ the equality holds (i.e. $|E| = \lceil \frac{k+2}{2} |V| \rceil + c(k)$). We prove that for such a strongly minimal k-vertex globally rigid graph $c(k) \geq 0$ always holds.

First suppose k is even. Then, we will have $\delta_{avg} = k + 2 + \frac{2c(k)}{|V|} \geq k+2$, which implies $c(k) \geq 0$.

If k is odd, then $|E| = \frac{k+1}{2} |V| + \left\lceil \frac{|V|}{2} \right\rceil + c(k)$ which gives $|E| < \frac{k+1}{2} |V| + c(k) + \frac{|V|}{2} + 1$. Therefore, $\delta_{avg} < k + 1 + \frac{2c(k)}{|V|} + 1 + \frac{2}{|V|}$ holds. This implies $k+2 < k+2+\frac{2c(k)}{|V|} + \frac{2}{|V|}$ which gives $-1 < c(k)$ and so $c(k) \geq 0$ holds. ■

Note that the underbound of Theorem 12 is achieved for $k = 2$, see Theorem 7. Below, we show it is also achievable for $k = 3$.

3.3 Strongly Minimal 3-vertex Global Rigidity in 2D

In this section we provide a class of graphs that are strongly minimal 3-vertex globally rigid. According to the important result of Theorem 12, for the case of 3-vertex global rigidity, the number of edges must satisfy the inequality $|E| \geq \lceil \frac{5}{2} |V| \rceil + c$. In order to characterize a strongly minimal 3-vertex globally rigid graphs, we conjecture that the number of edges has the form $|E| = \lceil \frac{5}{2} |V| \rceil + c$ and then try to find a suitable value for c. We also assume that the number of vertices is at least 6. Since we are seeking graphs with a minimum number of edges, it is reasonable to keep c at its minimum (0) and try to find a graph which has the desired property.

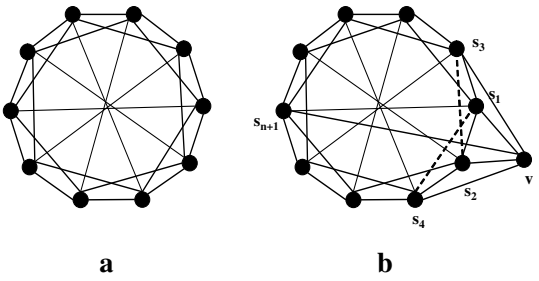


Fig. 4. (a) A Strongly Minimal 3-Vertex Globally Rigid graph of size 10 (R_{10}). (b) A Strongly Minimal 3-vertex globally rigid graph of size 11 (R_{11}).

If $|V|$ is even, i.e. $|V| = 2n$, the equation implies $|E| = \frac{5}{2}|V|$ and $\delta_{avg} = 5$. According to Lemma 2, $\delta(v) \geq 5$. Therefore, the graph we are seeking, is 5-regular, i.e. every vertex has degree 5. Fortunately, there is a class of graphs with these conditions that are 3-vertex globally rigid. An example of such a graph with $|V| = 10$ is shown in Figure 4a, and the structure carries over in an obvious way to graphs with $2n$ vertices for $n \geq 3$. The structure of such a graph, call it R_{2n}, is as follows:

The graph is formed from a C_{2n} cycle $[s_1, s_2, ...s_{2n}]$, to which are added the edges $s_1s_3, s_2s_4, s_3s_5, ...s_{2n-1}s_1, s_{2n}s_2$ (forming 2 cycles of size n denoted by $C_n^{(1)}$and $C_n^{(2)}$, respectively) and the edge set D of edges $s_is_{i+n}, i = 1..n$ (called diagonals). Note that the cycles $C_n^{(1)}$and $C_n^{(2)}$ are disjoint.

It needs to be proved that the proposed graph is 3-vertex globally rigid. The proof has two steps: first, showing the graph is 5-connected, and second, after removal of any 2 vertices, showing the result is redundantly rigid. Proofs are omitted due to space limitation.

Theorem 13. The graph R_{2n} on 6 or more vertices, depicted in Figure 4a, is 5-connected.

Theorem 14. The graph R_{2n} on 6 or more vertices, is 3-vertex globally rigid.

To obtain a class of strongly minimal 3-vertex globally rigid with an odd number of vertices, i.e. $|V| = 2n+1$, we define an operation, termed *2-extension*, over R_{2n} graphs to increase $|V|$ by 1. We show that this operation preserves 3-vertex global rigidity. The sketch of *2-extension* is as follows:

Assume vertices $s_i, i = 1..4$ to be 4 consecutive vertices in the C_{2n} cycle as depicted in Figure 4b. The diagonal neighbor of s_1 is called s_{n+1}. The operation consists of adding a new vertex, say v, connecting it to $s_i, i = 1..4, (n+1)$ and removing (s_1, s_4) and (s_2, s_3).

Let R_{2n+1} be the graph obtained from R_{2n} by applying the *2-extension operation* (Figure 4b). Evidently,

$|E(R_{2n+1})| = |E(R_{2n})| + 3 = \frac{5}{2}(2n) + 3 = \frac{5}{2}(2n+1) + \frac{1}{2}$ holds which is consistent with the edge count condition ($|E(R_{2n+1})| = \left\lceil \frac{5}{2} |V(R_{2n+1})| \right\rceil = \frac{5}{2}(|V(R_{2n+1})| - 1) + \frac{1}{2}$). Hence, R_{2n+1} has the minimum number of edges and by proving that R_{2n+1} is 3-vertex globally rigid, it follows that it is strongly minimal 3-vertex globally rigid.

Theorem 15. The 2-extension operation applied to R_{2n} preserves 3-vertex global rigidity.

Proof. The proof is similar to the proof of Theorem 14. The method is to consider all possible choices of U, $U \in V(R_{2n+1})$ and $|U| = 2$, and show that $R_{2n+1} \backslash U$ is globally rigid. We omit the proof due to the space limitation. ■

3.4 Comparison of Results

Table 1 shows the result of comparing the structures suggested for 3-vertex globally rigid graphs in Section 3.1 (8-connected) and Section 3.2 (strongly minimal). It is easy to see that the strongly minimal 3-vertex globally rigid structure is at least $\frac{8n-5n}{8n} = \frac{3}{8} = 37.5\%$ more efficient to localize the network (in terms of the minimum number of required distance measurements to achieve the specified tolerance to vertex loss), than the structure suggested based on connectivity. Of course, there will be other redundancies in the larger network which are not present in the strongly minimal network.

Table 1. Comparison of the minimum number of required distance measurements for the suggested structures with n vertices

Structure	Min # of Dist. Measurements
8-Connectivity	$8n$
S.M. 3-V-G Rigid	$5n$

4. CONCLUSION

In this paper we studied the structure of localizable sensor networks which are tolerant against the loss of multiple nodes (the network remains localizable after the loss of multiple nodes). This is done by introducing the notion of k-vertex globally rigid graphs in which after removal of any set of up to $k-1$ vertices, the resulting network still remains globally rigid (localizable). For $2D$ networks, we showed that a graph is k-vertex globally rigid if it is $(k+5)$-connected. For networks modeled by a random geometric graph, this reduction may enable us propose a critical transmission radius r for the nodes, above which the network becomes k-vertex globally rigid. Furthermore, By providing a lower bound on the edge count of k-vertex globally rigid graph in terms of the vertex count, we also proposed a class of 3-vertex globally rigid graphs with minimum edge count. Comparisons showed a considerable improvement achieved by this class over the 8-connectivity condition, in terms of the required number of distance measurements. This suggests that, there is certainly a benefit in studying the structure of strongly minimal networks for a given k (and for $k > 3$ this remains to be done). Indeed, the full characterization of (strongly minimal) k-vertex globally rigid graphs in addition to efficient algorithms to test this property on a given graph are still open problems that constitute our future work.

REFERENCES

Aspnes, J., Eren, T., Goldenberg, D., Morse, A., Whiteley, W., Yang, Y., Anderson, B., and Belhumeur, P. (2006). A theory of network localization. *IEEE Transactions on Mobile Computing*, 5(12), 1663–1678.

Connelly, R. (2005). Generic global rigidity. *Discrete and Computational Geometry*, 33(4), 549–563.

Graver, J., Servatius, B., and Servatius, H. (1993). *Combinatorial rigidity*. American Mathematical Society.

Jackson, B. and Jordan, T. (2005). Connected rigidity matroids and unique realizations of graphs. *Journal of Combinatorial Theory, Series B*, 94, 1–29.

Jackson, B. and Jordán, T. (2009). A sufficient connectivity condition for generic rigidity in the plane. *Discrete Applied Mathematics*, 157(8), 1965–1968.

Jordan, T. and Szabadka, Z. (2009). Operations preserving the global rigidity of graphs and frameworks in the plane. *Computational Geometry*, 42(6-7), 511–521.

Laman, G. (1970). On graphs and rigidity of plane skeletal structures. *Journal of Engineering Mathematics*, 4, 331–340.

Lovasz, L. and Yemini, Y. (1982). On generic rigidity in the plane. *SIAM Journal on Algebraic and Discrete Methods*, 3, 91.

Nagamochi, H. and Ibaraki, T. (2008). *Algorithmic aspects of graph connectivity*. Cambridge University Press New York, NY, USA.

Servatius, B. (1989). Birigidity in the plane. *SIAM Journal on Discrete Mathematics*, 2(4), 582–599.

Summers, T., Yu, C., and Anderson, B.D.O. (2008). Robustness to agent loss in vehicle formations and sensor networks. In *Proceedings of the 47th IEEE Conference on Decision and Control*, 1193–1199.

Wan, P. and Yi, C. (2004). Asymptotic critical transmission radius and critical neighbor number for k-connectivity in wireless ad hoc networks. In *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, 1–8. ACM.

Yu, C. and Anderson, B.D.O. (2009). Development of redundant rigidity theory for formation control. *International Journal of Robust and Nonlinear Control*, 19(13), 1427–1446.

Yu, C. and Anderson, B. (2008). Agent and Link Redundancy for Autonomous Formations. In *Proceedings of the 17th IFAC World Congress*, 6584–6589. IFAC.

Yu, C., Dusgupta, S., and Anderson, B.D.O. (2010). Robustness to agent loss in vehicle formations and sensor networks. Submitted for publication.

49th IEEE Conference on Decision and Control
December 15-17, 2010
Hilton Atlanta Hotel, Atlanta, GA, USA

Network Localizability with Link or Node Losses

Changbin Yu Soura Dasgupta Brian D. O. Anderson

Abstract— **The ability to localize a sensor network is important for its deployment. A theoretical result exists defining necessary and sufficient conditions for network unique localizability (for inter-sensor range-based localization); it has its roots in Graph Rigidity Theory where sensors and links/measurements are modelled as vertices and edges of a graph, respectively. However, critical missions do require a level of robustness for localizability, ensuring that localizability is retained in the event of link (edge) losses and/or sensor (vertex) losses. This work characterizes this robustness through a novel notion of redundant localizability, which is backed by redundant rigidity. Analogously to two well-known types of result for rigidity characterization, similar results are developed for edge redundant rigidity; they are supplemented by rather fewer results dealing with vertex redundant rigidity. These preliminary results may shed a light for any further study of redundant localizability.**

Index Terms— **Rigidity Theory, Network Localizability, Sensor Network Localization**

I. Introduction

Obtaining location information is a fundamental task in a sensor network, since otherwise the sensed data will become much less valuable. The location of every sensor node, if not already known from the deployment of the network or directly from GPS, can only be determined from a process based on measurements, the network structure and partially known location information. This process is referred to as localization, and the network property that governs the feasibility of localizing the entire network given the measurement is called localizability.

Most localization algorithms are based on range-measurements, or measurements that can be converted to inter-node distance measurements.

Aspens et al [2] formally prove that for a 2D localization problem, a necessary and sufficient condition for localizability given inter-node distance measurements is that a network graph must be globally rigid (the concept is reviewed below) and at least three noncollinear sensors must have known location information (hence they are anchors).

Having just global rigidity may be unrealistic in practical scenarios, as not only can the localization algorithms demand very high computational complexity, but also because the global rigidity property (hence localizability) can easily be lost if some node or measurement becomes unavailable, manifesting some type of node/link failure in the network.

C. Yu is with the Australian National University, Canberra ACT 2600, Australia. (brad.yu@anu.edu.au, Tel:+61-2-61258670, Fax: +61-2-61258660.)

S. Dasgupta is with Department of Electrical and Computer Engineering, University of Iowa (Dasgupta@engineering.uiowa.edu)

B.D.O. Anderson is with the Australian National University and NICTA Ltd. (Brian.Anderson@anu.edu.au)

The first problem has attracted great interest among researchers. Anderson et al [1] discussed various graphical properties of easily localizable sensor networks. A main result is that by doubling or tripling the sensing radius, the new network graph acquires special properties that provide guarantees on the computational complexity, sometimes linear in the number of sensors. In [14], limitations of classical trilateration algorithms are analyzed, proving their insufficiency in even recognizing the localizability of a graph. A novel localization method generalizing trilateration is proposed based on aggregated knowledge of the subnetwork formed by all 1-hop neighbors of each node and the node itself; roughly speaking, the induced subgraph of the neighbors and the node is actually a wheel graph that is globally rigid.

The second problem, viz. ensuring tolerance of node or link failure, has been rarely addressed. Recent study of redundant rigidity [13] has started to shed light on a new direction, using a graph theoretical approach. It raises the question, pursued in depth in this paper: *what are the conditions for preserving localizability in the event of loss of up to p link length measurements and q nodes (together with all their associated distance measurements)?*

This work studies the level of redundancy that can be built into localizability, such that, the network is guaranteed to be localizable when the loss of nodes and/or links is allowed, up to a certain maximum number in each case. As the result is more of a fundamental characterization than a practical algorithm, the typical problems in practical localization, such as dealing with noisy measurements, characterization of RMS error and bias in position estimates, error propagation, and computational complexity, are not considered and remain for the future.

The rest of this paper is organized as follows. The preliminaries, especially relevant results of graph rigidity theory, are introduced in Section II. The connections between redundant rigidity, redundant connectivity, global rigidity and redundant localizability are established in Section III. Based on these, Section IV discusses the relationship between vertex and edge redundancy. Two characterizations of $(1, p)$-rigidity, a concept associated with the loss of links, are then given in Section V and Section VI, respectively. Section VII sets out the relationship between redundant connectivity and $(1, p)$-rigidity on the one hand and redundant global rigidity and localizability on the other. Finally, the conclusions and proposals for future work are presented. Due to space limitation, all proofs are omitted and can be found in an extended version by a request to the first coauthor.

978-1-4244-7744-9/10/$26.00 ©2010 IEEE

II. PRELIMINARIES

We will model a network as an undirected graph $G = (V, E)$ where the nodes are elements of the vertex set V, and an edge $e_{ij} \in E$ if the distance between nodes i and j is available. In the sequel for $X \subset V$, $i_G(X)$ will be the number of edges in the graph induced by the vertex set X in G. Further, with $V_i \subset V$, the edge set in the graph induced by the vertex set V_i in G will be denoted by $E_G(V_i)$. Similarly, with $E_i \subset E$, the vertex set in the graph induced by the edge set E_i in G will be denoted by $V_G(E_i)$.

As noted in the introduction a result due to [2] proves that a two-dimensional network $G = (V, E)$ is localizable iff it is globally rigid and has at least three noncollinear anchor nodes. Thus in section II-A we provide some preliminary information about rigidity and global rigidity.

In subsection II-B we provide two alternative though, equivalent characterizations of rigidity. The first, a celebrated result due to Laman, [10] involves edge counts on subgraphs induced by components of vertex partitions. The second, primarily due to Lovasz and Yemeni [11] involves vertex counts on subgraphs induced by components of edge partitions.

Subsequently, in Section III we provide pertinent redundancy definitions and an important linkage between redundant rigidity, redundant connectivity and global rigidity, that sets up the remainder of the paper.

A. Rigidity and Global Rigidity

In this subsection, we recall briefly the notions and some properties of (minimal) rigidity and global rigidity. Our description will be largely based on graphs that model the network under consideration. Formal definitions of graph rigidity and global rigidity, which involve the theory of graph representations, can be found in, e.g., [4], [5], [12].

A graph is *rigid* if the only edge-length preserving smooth motions in a generic network[1] modelled by the graph are translation and rotation, thus only result in congruent graphs. A graph is *minimally rigid* if it is rigid and no single edge can be removed without losing rigidity.

Close inspection of the rigidity definition reveals that the consideration is made for smooth motions. Under discontinuous transformation of a network corresponding to a rigid graph, it is possible to have a non-congruent network that preserves the edge-lengths due to flip and/or flex ambiguity.

To eliminate these ambiguities and to uniquely determine the relative positions of any vertex given the set of distance measurements (corresponding to the edge-length set), a stronger definition is desirable. A graph is *globally rigid* if any two realizations of a network (with prescribed edge lengths and modelled by the graph) of the one edge-length set are congruent, i.e., differ at most by translation, rotation or reflection. We refer the reader to [4], [8] for source material on global rigidity.

[1]The graph theory literature generally uses a different term than network in discussing rigidity, the word framework being commonly employed. Because of this paper's connection with sensor networks, we retain the word network below.

B. Rigidity Characterization

We provide now two different sets of necessary and sufficient conditions for a graph to be rigid. The first due to Laman, [10] involves edge counts on graphs induced by vertex subsets.

Theorem 2.1: [Laman's Theorem] A graph $G = (V, E)$ is minimally rigid iff $|E| = 2|V| - 3$ and for all $X \subset V$, $i_G(X) \leq 2|X| - 3$. The graph is rigid iff there is an $E' \subset E$, such that $G' = (V, E')$ is minimally rigid.

In the sequel, we will also need the following result that presents Henneberg Operations for growing a rigid graph [7].

Theorem 2.2: [Vertex Addition] Consider a graph $G = (V, E)$, a node $k \notin V$ and two edges e_{ik} and e_{jk}, with i and j elements of V. Then G is (minimally) rigid iff $(V \bigcup k, E \bigcup \{e_{ik}, e_{jk}\}$ is (minimally) rigid.

Theorem 2.3: [Edge Splitting] Consider a graph $G = (V, E)$, a node $t \notin V$ and three edges e_{ti} e_{tj} and e_{tk}, with i j and k elements of V and e_{ij} element of E. Then G is (minimally) rigid if $(V \bigcup t, E \backslash e_{ij} \bigcup \{e_{ti}, e_{tj}, e_{tk}\}$ is (minimally) rigid.

There are efficient algorithms for checking the edge count condition underlying Laman's theorem [9].

The second necessary and sufficient condition due to Yemeni and Lovasz [11], involves vertex counts on graphs induced by edge subsets. Though less succinctly phrased, it may be easier to check than Laman's edge count result, and references to algorithms for checking the condition are provided in [11]. Jackson and Jordan [8] provide an elegant interpretation of this result using the theory of matroids.

Definition 2.1: Consider a graph $G = (V, E)$ and $E_i \subset E$. A set of subgraphs $P = \{G_i = (V_G(E_i), E_i)\}$ is termed an m-Admissible Decomposition (AD) of G if the following hold.

(i) $|E_i| > 0$.
(ii) $\bigcup_i E_i = E$.
(iii) P has at least m elements.

If $m = 1$, then P is simply called an AD of G.

Next we define the index of an m-AD that reflects an associated vertex count.

Definition 2.2: Consider a graph $G = (V, E)$ and $E_i \subset E$. We call $r(P)$ the index of an m-AD $P = \{G_i = (V_G(E_i), E_i)\}_{i=1}^{k}$ of G where $k \geq m$ is defined as

$$r(P) = \sum_{i=1}^{k} (2|V_G(E_i)| - 3)\,. \tag{II.1}$$

To state the result we need one more definition.

Definition 2.3: Consider a graph $G = (V, E)$ and $E_i \subset E$. For a given m, an m-AD $P = \{G_i = (V_G(E_i), E_i)\}_{i=1}^{k}$ of G, $k \geq m$ is termed m-minimizing if $r(P)$ is the smallest among the indices of all possible m-AD's of G. Such an index will be denoted as $r_m(G)$.

Now we provide the promised second characterization of rigidity that follows from results in [8], [11] and [6].

Theorem 2.4: A graph G is rigid iff $r_1(G) = 2|V| - 3$. Further, the edge sets E_i underlying any minimizing admissible decomposition are disjoint.

Observe one cannot say that $r_1(G) = 2|V| - 3$ implies *minimal* rigidity. Indeed consider the graph $G = K_4$. Then with $P = \{G\}$, $r(G) = 2|V| - 3$, even though G is not minimally rigid.

We conclude this section with an associated definition from [8] and two results, one from [8], that will assist us in later development.

Definition 2.4: Consider a graph $G = (V, E)$. Suppose S is a non-empty subset of E, and H is the subgraph induced by the edge set S. Then S is an independent subset of E if $i_H(X) \leq 2|X| - 3$ for all $X \subset V_G(S)$, as long as $|X| \geq 2$. The null set is also an independent subset of E.

The following result is a direct consequence of Laman's theorem.

Lemma 2.1: Suppose a graph $G = (V, E)$ is minimally rigid. Then for any $S \subset E$, S is a maximally independent subset of S in the subgraph of G, induced by S.

The final result in this section is a translation of Lemma 2.4 from [8].

Lemma 2.2: Consider a graph $G = (V, E)$ with $|E| \geq 1$. Suppose $S \subset E$ is a maximally independent subset of E. Then $r_1(G) = |S|$.

III. Redundant Global Rigidity

Generally, there are two types of redundancy: those involving loss of edges, and those involving the loss of vertices. As a generalization we work here with *mixed redundancy*. We begin with definitions of mixed redundant rigidity and mixed redundant connectivity.

Definition 3.1: A graph $G = (V, E)$ is (q, p)-rigid if for all $0 \leq l \leq q - 1$ and $0 \leq k \leq p - 1$, the induced subgraph obtained by removing any l vertices and k edges is rigid.

In particular, a $(q, 1)$-rigid graph is what has been defined in [13] as *q-vertex rigid*, or simply q-rigid. Similarly a $(1, p)$-rigid graph is known as *p-edge rigid*. There is also a comparable definition for redundant connectivity.

Definition 3.2: A graph $G = (V, E)$ is (q, p)-connected if for all $0 \leq l \leq q - 1$ and $0 \leq k \leq p - 1$, the induced subgraph obtained by removing any l vertices and k edges is connected.

These definitions bring us to our first new result, aimed at tightening them. Specifically, we observe below that these definitions are stronger than is needed. Indeed it is clear from Laman's theorem that if a graph is not rigid then the removal of a further edge cannot make it rigid. The same fact applies to connectivity. Thus it is possible to alter the definition of $(1, p)$-rigidity/connectivity to simply require that the graph remain rigid/connected after the removal of any $p-1$ edges, as opposed to *up to p-1 edges* as required by the foregoing definitions.

On the other hand, it is entirely possible for a nonrigid or a disconnected graph to gain rigidity or connectivity after losing vertices. Thus consider a rigid graph (V, E). Suppose $v \notin V$ and let e be an edge connecting v to a vertex in V. Then it is easy to see that the graph $(V \bigcup \{v\}, E \bigcup \{e\})$ is not rigid.

Similarly if the graph (V, E) is connected, the graph $(V \bigcup \{v\}, E)$ is not connected as the vertex v is isolated in this augmented graph.

We now assert that in fact, barring small graphs, even with q-vertex rigidity it suffices to check if rigidity is retained after deleting any $q - 1$-vertices.

Theorem 3.1: A graph $G = (V, E)$, with $|V| > q + 1$ is (q, p)-rigid iff it is rigid after removing any $q - 1$ vertices and $p - 1$ edges.

A similar result obtains for redundant connectivity.

Theorem 3.2: A graph $G = (V, E)$, with $|V| > q$ is (q, p)-connected iff it is connected after removing any $q-1$ vertices and $p - 1$ edges.

We now recount a result from [8] that ties connectivity and rigidity to global rigidity.

Theorem 3.3: A graph $G = (V, E)$ is globally rigid iff it is $(1, 2)$-rigid and $(3, 1)$-connected.

We now define *redundant global rigidity*, which as noted in section II is equivalent to redundant localizability given three or more noncollinear anchors.

Definition 3.3: A graph $G = (V, E)$ is (q, p)-globally rigid if it is globally rigid and for all $0 \leq l \leq q - 1$ and $0 \leq k \leq p - 1$, the induced subgraph obtained by removing any l vertices and k edges is globally rigid.

In view of theorems 3.1-3.3 one has the following result.

Theorem 3.4: A graph $G = (V, E)$, with $|V| > q + 1$ is (q, p)-globally rigid iff it is rigid after removing any $q - 1$ vertices and p edges and is connected after removing any $q + 1$ vertices and $p - 1$ edges.

Thus redundant global rigidity has two components: Redundant rigidity and redundant connectivity.

Finally, we come to redundant localizability.

Definition 3.4: A sensor network with a given set of inter-node distance measurements is (q, p) redundantly localizable if it remains localizable after removing any $q - 1$ nodes and any $p - 1$ edges.

Since a network is localizable if and only if it is globally rigid and has three or more noncollinear anchors, the previous theorem immediately yields:

Theorem 3.5: A two-dimensional sensor network with a given set of anchors and inter-node distance measurements and at least $q + 2$ nodes is (q, p) redundantly localizable if and only if its associated graph is rigid after removing any $q-1$ vertices and p edges, connected after removing any $q+2$ vertices and $p - 1$ edges, and the network retains three or more noncollinear anchors after removing any $q-1$ vertices.

The rest of the paper is largely concerned with the characterization of the redundant rigidity.

IV. Connections between vertex and edge redundancy

In this section we summarize relationships between vertex and edge redundancy. The uniform message is that vertex redundancy conditions are stronger than their edge counterparts.

First we present a result concerning redundant rigidity.

Theorem 4.1: A graph $G = (V, E)$, with $|V| \geq p + q + 3$ that is $(q + s, p)$-rigid for some $s > 0$, is $(q, p + s)$-rigid.

Next we address connectivity.

Theorem 4.2: A (q, p)-connected graph is $(q - s, p + s)$-connected for all $0 \leq s < q$.

We conclude this section by the following theorem that directly follows from Theorems 4.1 and 4.2, together with the characterization of redundnant global rigidity in terms of redundant rigidity and connectivity of Theorem 3.4.

Theorem 4.3: A graph $G = (V, E)$, with $|V| \geq p + q + 3$ that is $(q + s, p)$-globally rigid for some $s > 0$, is $(q, p + s)$-globally rigid.

V. Redundant Edge Rigidity: The Laman Approach

This section focuses on seeking a *Laman type* necessary and sufficient condition for $(1, p)$-rigidity. First we provide a natural definition of *minimal (1,p)-rigidity* generalized from the standard notion of (nonredundant) minimal rigidity, i.e. when $p = 1$.

Definition 5.1: A graph $G = (V, E)$ is minimally $(1, p)$-rigid if it is $(1, p)$-rigid and loses rigidity after the removal of any set of p edges.

It is not immediately clear that minimally $(1, p)$-rigid graphs actually exist for $p > 1$. It might be that a $(1, p)$-rigid graph necessarily has some sets of p edges whose removal renders the graph nonrigid, and some sets of p edges whose removal leaves the graph rigid. This is indeed the case for $p > 2$, see Lemma 5.1 below (though we do not prove this here); however for $p = 2$, the definition is meaningful, [8].

We first provide a necessary and sufficient condition for minimal $(1, 2)$-rigidity [8].

Theorem 5.1: A graph $G = (V, E)$ is minimally $(1, 2)$-rigid if and only if both the conditions below hold:

(a) $|E| = 2|V| - 2$.

(b) For any $X \subset V$, $1 < |X| < |V|$, $i_G(X) \leq 2|X| - 3$.

This result is in fact quite powerful, in that beyond a precise edge count, all it requires is that Laman's condition be satisfied by the graph induced by any proper subset of V.

Rigidity, as opposed to minimal rigidity, is characterized through Laman's theorem using the property that for a rigid graph $G = (V, E)$ there must exist an $E' \subset E$, such that $G = (V, E')$ is minimally rigid. The example depicted in Fig. 1(a) shows however, that *this is not true in general* for $(1, p)$-rigidity when $p > 1$.

To analyze the graph $G = (V, E)$, depicted in Fig. 1(a) observe that every vertex in the first and last rows of this nine vertex graph has an edge to every vertex in the second row. Thus in all there are 18 edges in this graph. Obviously, as

$$18 > 2 \times 9 - 4 + 2,$$

this graph cannot be minimally $(1, 2)$-rigid. It is also clear that there is no $E' \subset E$ such that (V, E') is minimally $(1, 2)$-rigid. This is so as were such an E' to exist it, $E \backslash E'$ must have precisely two elements in it. It is clear that the removal of any edge from G leaves at least one vertex with degree equal to two. Then no matter what further edge is removed, the resulting subgraph cannot be $(1, 2)$-rigid.

We now assert that G is in fact $(1, 2)$-rigid. To see this first observe that the graph induced by the vertices $\{1, \cdots, 6\}$ and that induced by $\{4, \cdots, 9\}$ are each minimally rigid. Indeed the graph in Fig. 1(b) is minimally rigid. The graph induced by $\{1, \cdots, 6\}$ can be built from this by two successive edge-splitting operations. The minimal rigidity of the graph induced by $\{4, \cdots, 9\}$ similarly follows.

Also note that there is a complete symmetry between the edges, in that if we show that the graph obtained by removing the edge e_{14} is rigid, this implies that a graph obtained by removing any single edge is rigid. It is easy to see that if edge e_{14} is removed, the graph remains rigid. Thus indeed the graph in Fig. 1 is $(1, 2)$-rigid.

The above observation shows that a $(1, p)$-rigid graph may not contain a minimally $(1, p)$-rigid graph. However in addition, the definition of minimally $(1, p)$-rigid graph is restricted to $p \leq 2$ only. Seeking an alternative definition remains open problem.

Lemma 5.1: For every integer $p > 2$, there is no minimally $(1, p)$-rigid graph satisfying Definition 5.1.

VI. Redundant Edge Rigidity: The Lovasz-Yemeni Approach

In this section, we characterize $(1, p)$-rigidity using the Lovasz-Yemeni approach. As in section V we will state the result attending first to minimally $(1, 2)$-rigid graphs.

Lemma 6.1: Let $G = (V, E)$ be a minimally $(1, 2)$-rigid graph. Then $r_2(G) = 2|V| - 2$, where $r_2(G)$ is defined in Definition 2.3.

We next state a result that shows that $r_p(G) \geq 2|V| - 4 + p$ is a sufficient condition for $(1, p)$-rigidity.

Theorem 6.1: Let $G = (V, E)$ be a graph with $r_p(G) \geq 2|V| - 4 + p$. Then G is $(1, p)$-rigid.

Combining these two results we immediately obtain the following necessary and sufficient condition for minimal $(1, 2)$-rigidity.

Theorem 6.2: A graph $G = (V, E)$ is minimally $(1, 2)$-rigid iff $r_2(G) \geq |E| = 2|V| - 2$.

The question remains whether $r_p(G) = 2|V| + p - 4$ is a necessary condition for *nonminimal* $(1, p)$-rigidity (assuming that $|E| \neq 2|V| + p - 4$). The graph $G = (V, E)$ in Fig. 2 serves as a counter example.

Consider the three edge partitions: $E_1 = E_G(\{1, 2, 4, 5\})$, $E_2 = E_G(\{3, 7, 4, 8\})$ and $E_3 = E_G(\{8, 9, 6, 5\})$. $P = \{(V_G(E_i), E_i)\}_{1=1}^{3}$ is a 2-AD for G. In this case:

$$r(P) = 3 \times (8 - 3) = 15 < 2 \times 9 - 2 = 16.$$

The graph is clearly rigid after removing any single edge other than from the set e_{45}, e_{48}, e_{58}. Now without sacrificing generality remove the edge e_{45}. Observe it can be reconstructed by an edge-splitting operation on $(V \setminus \{2\}, E_G(V \setminus \{2\}))$. Thus this graph is $(1, 2)$-rigid.

To this point we have shown that both Lovasz-Yemeni and Laman type necessary and sufficient conditions exist for minimal $(1, p)$-rigidity ($p = 2$), the Lovasz-Yemeni

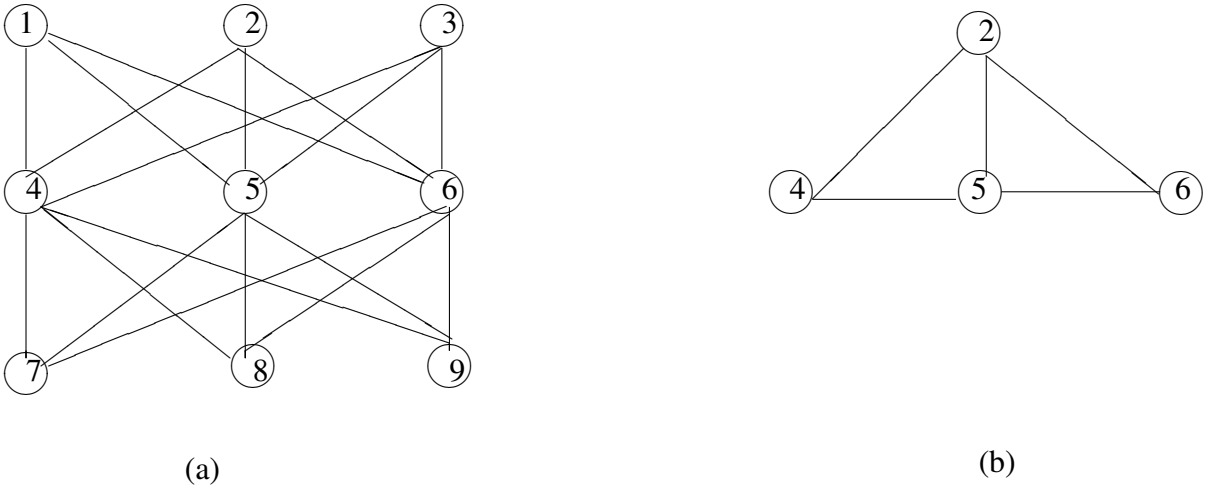


Fig. 1. (a) A $(1,2)$-rigid graph that does not contain a minimally $(1,2)$-rigid subgraph; (b) A graph used in proving claim for graph of (a).

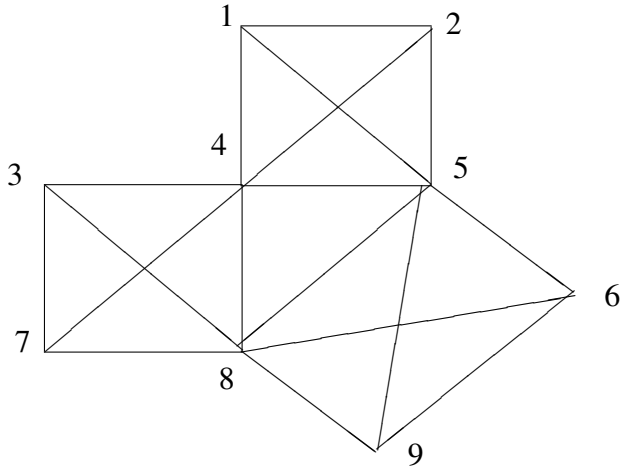


Fig. 2. A $(1,2)$-rigid graph with $r_2(P) < 2|V| - 2$.

type sufficient condition also characterize nonminimal $(1,p)$-rigidity for all positive integer p.

Recall however, that the direct subject of investigation in this paper is redundant global rigidity, which beyond redundant rigidity also requires redundant connectivity. It is instructive to note that the example of Fig. 1 is *(3,1)-connected, though that of Fig. 2 is not*. It is also possible to show that the graph in Fig. 1 does satisfy the corresponding Lovasz-Yemeni sufficient condition for $(1,2)$-rigidity. Indeed the next section demonstrates that a $(1,p)$-rigid graph that has an additional redundant connectivity property must have $r_p(G) \geq 2|V| - 4 + p$.

Clearly, having an efficient algorithm for computing r_p is important. While we fail to include such an algorithm in this paper, we would like to conjecture that it exists: Lovasz and Yemeni [11] claimed that $r_1(G)$ is easy to compute, and by definition finding $r_p(G)$, $p > 1$ involves computation of sums of vertex cardinalities involving a set of subgraphs that is also a subset of those needed to compute $r_1(G)$.

VII. Lovasz-Yemeni type condition for redundant rigidity with redundant connectivity

This section considers levels of connectivity needed for a $(1,p)$-rigid graph to satisfy $r_p(G) \geq 2|V| - 4 + p$. At the end of this section we will tie the results here to redundant global rigidity.

We have the following result connecting Yemeni-Lovasz type conditions to $(1,p)$-global rigidity for $p \leq 4$.

Theorem 7.1: Suppose, for integer p with $2 \leq p \leq 4$, $G = (V,E)$ is $(3,p-1)$-connected and $(1,p)$-rigid. Then $r_p(G) \geq 2|V| - 4 + p$.

There remains the question whether the result of Theorem 7.1 will hold for $p > 4$. We now present a counterexample for $p = 5$.

Example 7.1: The graph in question, $G = (V,E)$ has $|V| = 33$, and has a vertex cover $V_1, \cdots, V_6$, such that each $|V_i| = 7$, each subgraph $(V_i, E_G(V_i))$ is a K_7 graph, and the sets $E_i = E_G(V_i)$ consitute a partition of E. Further, defining

$$X_i = V_i \cap (\cup_{j \neq i} V_j)\,, \tag{VII.2}$$

one has $X_1 = \{1,2,3\}$, $X_2 = \{1,4,5\}$, $X_3 = \{2,6,7\}$, $X_4 = \{3,8,9\}$, $X_5 = \{4,6,8\}$ and $X_6 = \{5,7,9\}$. In other words V_i comprises X_i together with four other vertices that are not in any other V_i. Thus there are 24 vertices each belonging to just one V_i, and 9 which are in exactly two.

To see that this choice of X_i is consistent with the fact that the E_i partition E, observe for all $i \neq j$,

$$|X_i \cap X_j| \leq 1. \tag{VII.3}$$

Further vertices in each V_i connect to the others through the elements of X_i. Thus, indeed this defintion is consistent with the requirement that the E_i are disjoint.

In effect then G comprises six K_7 subgraphs, connected through the vertices in the X_i. Each V_i has three vertices through which it connects to other V_j, there being one vertex in common with each of three different V_j.

One can show that G is both $(3,4)$-connected and $(1,5)$-rigid. Since each subgraph $(V_i, E_G(V_i))$ is K_7, it also follows that $V_i = V_G(E_i)$. Thus as the E_i partition E, we have that

$$P = \{(V_i, E_G(V_i))\}_{i=1}^{6}$$

is a 6-AD of G. Hence,

$$r_5(G) \leq r(P) = 6(14-3) = 66 = 2|V| < 2|V| - 4 + 5,$$

and Theorem 7.1 is violated for $p = 5$.

Nonetheless we now show that for $p > 4$, a stronger redundant connectivity condition suffices for $r_p(G) \geq 2|V| - 4 + p$ to hold given $(1,p)$-rigidity.

Theorem 7.2: Suppose, for integer $p \geq 3$, $G = (V,E)$ is $(4, p-2)$-connected and $(1,p)$-rigid. Then $r_p(G) \geq 2|V| - 4 + p$.

As global rigidity is equivalent to the combined properties of $(3,1)$-connectivity and $(1,2)$-rigidity, in view of the two theorems in this section we have the following key result.

Theorem 7.3: A graph $G = (V,E)$ is $(1,p)$-globally rigid if it is $(3,p)$-connected and $r_{p+1}(G) \geq 2|V| - 3 + p$. For $1 \leq p \leq 3$ it is $(1,p)$-globally rigid only if $r_{p+1}(G) \geq 2|V| - 3 + p$. For $p \geq 4$ a $(4, p-2)$-connected $G = (V,E)$ is $(1,p)$-globally rigid only if $r_{p+1}(G) \geq 2|V| - 3 + p$.

Given the relation between global rigidity, anchor count and localizability, there is an obvious extension of this to a corresponding characterization of redundantly localizable two-dimensional sensor networks.

We comment that in a sensor network in which loss of nodes is contemplated, it may be that loss of anchor nodes can be ruled out on reliability or other grounds. What is needed for redundant localizability in addition to redundant global rigidity is that after loss of up to q-1 nodes, at least three noncollinear anchor nodes remain. The issue of which nodes fail does not enter the picture when the redundancy is all with respect to edge loss of course.

VIII. Conclusions and Future Work

This work presents preliminary results on edge-redundant global rigidity, which is essential to the problem of seeking a solution to guaranteing network localizability in the event of link losses. The particular notion that is studied in detail is $(1,p)$-rigidity, for which two different types of characterizations were obtained: Laman type and Lovasz-Yemeni type. For the latter, further connection was established between redundant edge rigidity and redundant connectivity. We also show that a seemingly obvious definition of minimal $(1,p)$ rigidity is meaningless for $p > 2$.

As discussed in the beginning of this paper, it is of great importance to also study redundant vertex rigidity, i.e. $(q,1)$-rigidity. Some discussion is provided in the text but its characterization remains largely open. Also, one will need computationally efficient algorithms to check for satisfaction of these conditions (e.g. computing $r_p(G)$) for redundant localizability of a network. Moreover, operations are required for one to construct, augment, merge and split such networks, while ensuring the level of redundant localizability remain unchanged. Results along the lines of [1] providing simple sufficient conditions for redundant global rigidity, largely in terms of local properties of graphs, would also be welcome. At some point too, random geometric graphs should be investigated.

Acknowledgement

This work is supported by USAF-AOARD-09-4136. C. Yu is supported by the Australian Research Council (ARC) through an Australian Postdoctoral Fellowship under DP-0877562. S. Dasgupta is supported by US NSF grants ECS-0622017, CCF-072902, and CCF-0830747. B.D.O. Anderson is supported by the ARC and National ICT Australia (NICTA).

References

[1] B. D. O. Anderson, P. N. Belhumeur, T. Eren, D. K. Goldenberg, A. S. Morse, W. Whiteley, and Y. R. Yang. Graphical properties of easily localizable sensor networks. *Wireless Networks*, 15(2):177–191, February 2009.

[2] J. Aspnes, T. Eren, D. Goldenberg, A. S. Morse, W. Whiteley, Y. Yang, B. D. O. Anderson, and P. Belhumeur. A theory of network localization. *IEEE Transactions on Mobile Computing*, 5(12):1663–1678, December 2006.

[3] B. Bolobas. *Modern Graph Theory*. Springer, 1998.

[4] R. Connelly. Generic global rigidity. *Discrete and Computational Geometry*, 33:549–563, 2005.

[5] T. Eren, B.D.O. Anderson, A.S. Morse, W. Whiteley, and P.N. Belhumeur. Operations on rigid formations of autonoumous agents. *Communications in Information and Systems*, 3(4):223–258, September 2004.

[6] J. E. Graver, B. Servatius, and H. Servatius. *Combinatorial Rigidity*. Graduate Studies in Mathematics. American Mathematical Society, 1993.

[7] L. Henneberg. In *Die graphische Statik der starren Systeme.*, Leipzig, 1911.

[8] B. Jackson and T. Jordan. Connected rigidity matroids and unique realizations of graphs. *Journal of Combinatorial Theory Series B*, 94:1–29, 2005.

[9] Donald J. Jacobs and Bruce Hendrickson. An algorithm for two-dimensional rigidity percolation: the pebble game. *J. Comput. Phys.*, 137(2):346–365, 1997.

[10] G. Laman. On graphs and rigidity of plane skeletal structures. *Journal of Engineering Mathematics*, 4(4):331–340, October 1970.

[11] L. Lovász and Y. Yemini. On generic rigidity in the plane. *SIAM Journal on Algebraic and Discrete Methods*, 3(1):91–98, 1982.

[12] W. Whiteley. Some matroids from discrete applied geometry. In J. E. Bonin, J. G. Oxley, and B. Servatius, editors, *Matroid Theory*, volume 197 of *Contemporary Mathematics*, pages 171–311. American Mathemtical Society, 1996.

[13] C. Yu and B.D.O. Anderson. Development of redundant rigidity theory for formation control. *International Journal of Robust and Nonlinear Control*, 19(13):1427–1446, 2009.

[14] Y. Liu Z. Yang and X. Li. Beyond trilateration: On the localizability of wireless ad-hoc networks. In *Proc. of the IEEE INFOCOM*, Rio de Janeiro, Brazil, 2009.

Redundant Localizability of Sensor Networks ⋆

Changbin Yu, Soura Dasgupta, Brian D.O. Anderson

Abstract

The ability to localize a sensor network is important for its deployment. A theoretical result exists defining necessary and sufficient conditions for network unique localizability (for inter-sensor range-based localization); it has its roots in Graph Rigidity Theory where sensors and links/measurements are modelled as vertices and edges of a graph, respectively. However, critical missions do require a level of robustness for localizability, ensuring that localizability is retained in the event of link (edge) losses and/or sensor (vertex) losses. This work characterizes this robustness through a novel notion of redundant localizability, which is backed by redundant rigidity. Analogously to two well-known types of result for rigidity characterization, similar results are developed for edge redundant rigidity; they are supplemented by rather fewer results dealing with vertex redundant rigidity. These results form a foundation for any further study of redundant localizability.

Key words:
Rigidity Theory, Network Localizability, Sensor Network Localization

1 Introduction

Obtaining location information is a fundamental task in a sensor network, since otherwise the sensed data will become much less valuable. The location of every sensor node, if not already known from the deployment of the network or directly from GPS, can only be determined from a process based on measurements, the network structure and partially known location information. This process is referred to as localization, and the network property that governs the feasibility of localizing the entire network given the measurement is called localizability.

Most localization algorithms are based on range-measurements, or measurements that can be converted to inter-node distance measurements.

Aspens et al [2] formally prove that for a 2D localization problem, a necessary and sufficient condition for localizability given inter-node distance measurements is that a network graph must be globally rigid (the concept is reviewed below) and at least three noncollinear sensors must have known location information (hence they are anchors).

Having just global rigidity may be unrealistic in practical scenarios, as not only can the localization algorithms demand very high computational complexity, but also because the global rigidity property (hence localizability) can easily be lost if some node or measurement becomes unavailable, manifesting some type of node/link failure in the network.

The first problem has attracted great interest among researchers. Anderson et al [1] discussed various graphical properties of easily localizable sensor networks. A main result is that by doubling or tripling the sensing radius, the new network graph acquires special properties that provide guarantees on the computational complexity, sometimes linear in the number of sensors. In [14], limitations of classical trilateration algorithms are analyzed, proving their insufficiency in even recognizing the localizability of a graph. A novel localization method generalizing trilateration is proposed based on aggregated knowledge of

⋆ A preliminary version of this paper has been published in the Proceedings of the 49th IEEE Conference on Decision and Control (CDC 2010), entitled 'Network Localizability with Link or Node Losses'. This work was supported by USAF-AOARD-10-4102. C. Yu is supported by the Australian Research Council (ARC) through a Queen Elizabeth II Fellowship under DP-110100538 and the Overseas" Expert Program of Shandong Province. S. Dasgupta is supported by US NSF grants ECS-0622017, CCF-072902, and CCF-0830747. B.D.O. Anderson is supported by the ARC and National ICT Australia (NICTA).

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

the subnetwork formed by all 1-hop neighbors of each node and the node itself; roughly speaking, the induced subgraph of the neighbors and the node is actually a wheel graph that is globally rigid.

The second problem, viz. ensuring tolerance of node or link failure, has been rarely addressed. Recent study of redundant rigidity [13] has started to shed light on a new direction, using a graph theoretical approach. It raises the question, pursued in depth in this paper: *what are the conditions for preserving localizability in the event of loss of up to p link length measurements and q nodes (together with all their associated distance measurements)?*

This work studies the level of redundancy that can be built into localizability, such that, the network is guaranteed to be localizable when the loss of nodes and/or links is allowed, up to a certain maximum number in each case. As the result is more of a fundamental characterization than a practical algorithm, the typical problems in practical localization, such as dealing with noisy measurements, characterization of RMS error and bias in position estimates, error propagation, and computational complexity, are not considered and remain for the future.

The rest of this paper is organized as follows. The preliminaries, especially relevant results of graph rigidity theory, are introduced in Section 2. The connections between redundant rigidity, redundant connectivity, global rigidity and redundant localizability are established in Section 3. Based on these, Section 4 discusses the relationship between vertex and edge redundancy. Two characterizations of minimal $(1,p)$-rigidity, a concept associated with the loss of links, are then given in Section 5 and Section 6, respectively. Section 7 establishes the relationship between redundant connectivity and $(1,p)$-rigidity on the one hand and redundant global rigidity and localizability on the other. Finally, the conclusions and future work are presented.

2 Preliminaries

We will model a network as an undirected graph $G = (V,E)$ where the nodes are elements of the vertex set V, and an edge $e_{ij} \in E$ if the distance between nodes i and j is available. In the sequel for $X \subset V$, $i_G(X)$ will be the number of edges in the graph induced by the vertex set X in G. Further, with $V_i \subset V$, the edge set in the graph induced by the vertex set V_i in G will be denoted by $E_G(V_i)$. Similarly, with $E_i \subset E$, the vertex set in the graph induced by the edge set E_i in G will be denoted by $V_G(E_i)$.

As noted in the introduction a result due to [2] proves that a two-dimensional network $G = (V,E)$ is localizable iff it is globally rigid and has at least three noncollinear anchor nodes. Thus in section 2.1 we provide some preliminary information about rigidity and global rigidity.

In subsection 2.2 we provide two alternative though, equivalent characterizations of rigidity. The first, a celebrated result due to Laman, [10] involves edge counts on subgraphs induced by components of vertex partitions. The second, primarily due to Lovasz and Yemeni [11] involves vertex counts on subgraphs induced by components of edge partitions.

Subsequently, in Section 3 we provide pertinent redundancy definitions and an important linkage between redundant rigidity, redundant connectivity and global rigidity, that sets up the remainder of the paper.

2.1 Rigidity and Global Rigidity

In this subsection, we recall briefly the notions and some properties of (minimal) rigidity and global rigidity. Our description will be largely based on graphs that model the network under consideration. Formal definitions of graph rigidity and global rigidity, which involve the theory of graph representations, can be found in, e.g., [4,5,12].

A graph is *rigid* if the only edge-length preserving smooth motions in a generic network [1] modelled by the graph are translation and rotation, thus only result in congruent graphs. A graph is *minimally rigid* if it is rigid and no single edge can be removed without losing rigidity.

Figure 1 shows several examples of two dimensional graphs, two of which are rigid and one of which is not rigid. In a network corresponding to non-rigid graph part of the network can 'flex' or move while preserving the edge-lengths, while the rest of the network stays still. The notion of rigidity conforms to one's normal intuition.

[1] The graph theory literature generally uses a different term than network in discussing rigidity, the word framework being commonly employed. Because of this paper's connection with sensor networks, we retain the word network below.

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

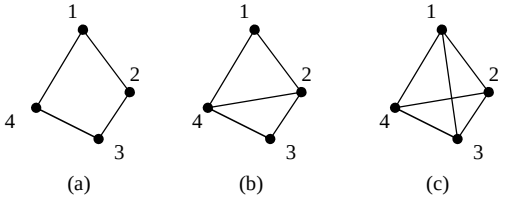


Fig. 1. 2D Realization of graphs that is (a) non-rigid, (b) (minimally) rigid, and (c) (globally) rigid.

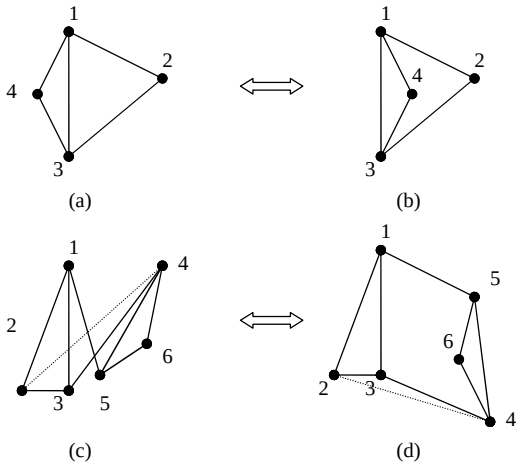


Fig. 2. Illustration of flip-ambiguity from (a) to (b); and flex-ambiguity from (c) to (d) . Lengths of all corresponding edges are the same.

Close inspection of the rigidity definition reveals that the consideration is made for smooth motions. Under discontinuous transformation of a network corresponding to a rigid graph, it is possible to have non-congruent network that preserves the edge-lengths due to flip and/or flex ambiguity. This is illustrated in Fig. 2.

To eliminate these ambiguities and to uniquely determine the relative positions of any vertex given the set of distance measurements (corresponding to the edge-length set), a stronger definition is desirable. A graph is *globally rigid* if any two realizations of a network (with prescribed edge lengths and modelled by the graph) of the one edge-length set are congruent, i.e., differ at most by translation, rotation or reflection. We refer the reader to [4, 8] for source material on global rigidity.

2.2 Rigidity Characterization

We provide now two different sets of necessary and sufficient conditions for a graph to be rigid. The first due to Laman, [10] involves edge counts on graphs induced by vertex subsets. We present this theorem below.

Theorem 2.1 *[Laman's Theorem] A graph $G=(V,E)$ is minimally rigid iff $|E|=2|V|-3$ and for all $X\subset V$, $i_G(X)\leqslant 2|X|-3$. The graph is rigid iff there is an $E'\subset E$, such that $G'=(V,E')$ is minimally rigid.*

In the sequel, we will also need the following result that presents Henneberg Operations for growing a rigid graph [7].

Theorem 2.2 *[Vertex Addition] Consider a graph $G=(V,E)$, a node $k\notin V$ and two edges e_{ik} and e_{jk}, with i and j elements of V. Then G is (minimally) rigid iff $(V\bigcup k, E\bigcup\{e_{ik},e_{jk}\}$ is (minimally) rigid.*

Theorem 2.3 *[Edge Splitting] Consider a graph $G=(V,E)$, a node $t\notin V$ and three edges e_{ti} e_{tj} and e_{tk}, with i j and k elements of V and e_{ij} element of E. Then G is (minimally) rigid iff $(V\bigcup t, E\backslash e_{ij}\bigcup\{e_{ti},e_{tj},e_{tk}\}$ is (minimally) rigid.*

There are efficient algorithms for checking the edge count condition underlying Laman's theorem [9].

The second necessary and sufficient condition due to Yemeni and Lovasz [11], involves vertex counts on graphs induced by edge subsets. Though less succinctly phrased, it may be easier to check than Laman's edge count result, and references to algorithms for checking the condition are provided in [11]. Jackson and Jordan [8] provide an elegant interpretation of this result using the theory of matroids. However, even though we rely heavily on results from [8], to avoid the need for a primer on matroids, we eschew matroid terminology, and use our own somewhat nonstandard but equivalent terminology.

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

Definition 2.1 *Consider a graph $G=(V,E)$ and $E_i \subset E$. A set of subgraphs $P=\{G_i=(V_G(E_i),E_i)\}$ is termed an m-Admissible Decomposition (AD) of G if the following hold.*

(i) $|E_i|>0$.
(ii) $\bigcup_i E_i = E$.
(iii) *P has at least m elements.*

If $m=1$, then P is simply called an AD of G.

Next we define the index of an m-AD that reflects an associated vertex count.

Definition 2.2 *Consider a graph $G=(V,E)$ and $E_i \subset E$. We call $r(P)$ the index of an m-AD $P=\{G_i=(V_G(E_i),E_i)\}_{i=1}^k$ of G where $k \geqslant m$ is defined as*

$$r(P)=\sum_{i=1}^{k}\left(2|V_G(E_i)|-3\right). \tag{2.1}$$

To state the result we need one more definition.

Definition 2.3 *Consider a graph $G=(V,E)$ and $E_i \subset E$. For a given m, an m-AD $P=\{G_i=(V_G(E_i),E_i)\}_{i=1}^k$ of G, $k \geqslant m$ is termed m-minimizing if $r(P)$ is the smallest among the indices of all possible m-AD's of G. Such an index will be denoted as $r_m(G)$.*

Now we provide the promised second characterization of rigidity that follows from results in [8], [11] and [6].

Theorem 2.4 *A graph G is rigid iff $r_1(G)=2|V|-3$. Further, the edge sets E_i underlying any minimizing admissible decomposition are disjoint.*

Observe one cannot say that $r_1(G)=2|V|-3$ implies *minimal* rigidity. Indeed consider the graph $G=K_4$. Then with $P=\{G\}$, $r(G)=2|V|-3$, even though G is not minimally rigid.

We conclude this section with an associated definition from [8] and two results, one from [8], that will assist us in later development.

Definition 2.4 *Consider a graph $G=(V,E)$. Suppose S is a non-empty subset of E, and H is the subgraph induced by the edge set S. Then S is an independent subset of E if $i_H(X) \leqslant 2|X|-3$ for all $X \subset V_G(S)$, as long as $|X| \geqslant 2$. The null set is also an independent subset of E.*

The following result is a direct consequence of Laman's theorem.

Lemma 2.1 *Suppose a graph $G=(V,E)$ is minimally rigid. Then for any $S \subset E$, S is a maximally independent subset of S in the subgraph of G, induced by S.*

The final result in this section is a translation of Lemma 2.4 from [8].

Lemma 2.2 *Consider a graph $G=(V,E)$ with $|E| \geqslant 1$. Suppose $S \subset E$ is a maximally independent subset of E. Then $r_1(G)=|S|$.*

3 Redundant Global Rigidity

Generally, there are two types of redundancy: those involving loss of edges, and those involving the loss of vertices. As a generalization we work here with *mixed redundancy*. We begin with definitions of mixed redundant rigidity and mixed redundant connectivity.

Definition 3.1 *A graph $G=(V,E)$ is (q,p)-rigid if for all $0 \leqslant l \leqslant q-1$ and $0 \leqslant k \leqslant p-1$, the induced subgraph obtained by removing any l vertices and k edges is rigid.*

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

In particular, a $(q,1)$-rigid graph is what has been defined in [13] as *q-vertex rigid*, or simply q-rigid. Similarly a $(1,p)$-rigid graph is known as *p-edge rigid*. There is also a comparable definition for redundant connectivity.

Definition 3.2 *A graph $G=(V,E)$ is (q,p)-connected if for all $0 \leqslant l \leqslant q-1$ and $0 \leqslant k \leqslant p-1$, the induced subgraph obtained by removing any l vertices and k edges is connected.*

These definitions bring us to our first new result, aimed at tightening them. Specifically, we show below that these definitions are stronger than is needed. Indeed it is clear from Laman's theorem that if a graph is not rigid then the removal of a further edge cannot make it rigid. The same fact applies to connectivity. Thus it is possible to alter the definition of $(1,p)$-rigidity/connectivity to simply require that the graph remain rigid/connected after the removal of any $p-1$ edges, as opposed to *up to p-1 edges* as required by the foregoing definitions.

On the other hand, it is entirely possible for a nonrigid or a disconnected graph to gain rigidity or connectivity after losing vertices. Thus consider a rigid graph (V,E). Suppose $v \notin V$ and let e be an edge connecting v to a vertex in V. Then it is easy to see that the graph $(V \bigcup \{v\}, E \bigcup \{e\})$ is not rigid.

Similarly if the graph (V,E) is connected, the graph $(V \bigcup \{v\}, E)$ is not connected as the vertex v is isolated in this augmented graph.

We now show that in fact, barring small graphs, even with q-vertex rigidity it suffices to check if rigidity is retained after deleting any $q-1$-vertices.

Theorem 3.1 *A graph $G=(V,E)$, with $|V| > q+1$ is (q,p)-rigid iff it is rigid after removing any $q-1$ vertices and $p-1$ edges.*

Proof: It is clear that a graph that is is (q,p)-rigid must be rigid after removing any $q-1$ vertices and $p-1$ edges. Now suppose the graph is rigid after removing any $q-1$ vertices and $p-1$ edges. To establish a contradiction, suppose the graph is not (q,p)-rigid. Observe if a graph is not rigid, then no graph with the same vertices and a subset of edges can be rigid. Thus there is a set of vertices V', with $0 \leqslant |V'| = l < q-1$ and a set of $p-1$ edges E_p, such that with $V'' = V \setminus V'$ the graph $G'' = (V'', E_G(V'') \setminus E_p)$ is not rigid but for some $j \in V''$ the graph $G'' = (V'' \setminus \{j\}, E_G(V'' \setminus \{j\}) \setminus E_p)$ is rigid. Then because of Theorem 2.2, j must have at most one edge in $G'' = (V'', E_G(V'') \setminus E_p)$. Now there are at least $q+1-l > 2$ vertices in $V'' \setminus \{j\}$. For any $\{k_1, \cdots, k_{q-1-l}\} \subset V'' \setminus \{j\}$, consider the set $V''' = V'' \setminus \{k_1, \cdots, k_{q-1-l}\}$. Observe $|V'''| = |V| - q + 1$. Thus by assumption $(V''', E_G(V''') \setminus E_p)$ is rigid. At the same time V''' has at least three elements, contains j, and yet j has at most one edge in V''', establishing a contradiction. ■

A similar result obtains for redundant connectivity.

Theorem 3.2 *A graph $G=(V,E)$, with $|V| > q$ is (q,p)-connected iff it is connected after removing any $q-1$ vertices and $p-1$ edges.*

Proof: It is clear that a graph that is (q,p)-connected must be connected after removing any $q-1$ vertices and $p-1$ edges. Now suppose the graph is connected after removing any $q-1$ vertices and $p-1$ edges. To establish a contradiction, suppose the graph is not (q,p)-connected. Observe if a graph is not connected, then no graph with the same vertices and a subset of edges can be connected. Thus there is a set of vertices V', with $0 \leqslant |V'| = l < q-1$ and a set of $p-1$ edges E_p, such that with $V'' = V \setminus V'$ the graph $G'' = (V'', E_G(V'') \setminus E_p)$ is not connected but for some $j \in V''$ the graph $G'' = (V'' \setminus \{j\}, E_G(V'' \setminus \{j\}) \setminus E_p)$ is connected. Then j must be isolated in $G'' = (V'', E_G(V'') \setminus E_p)$. Now there are at least $q+1-l > 1$ vertices in $V'' \setminus \{j\}$. For any $\{k_1, \cdots, k_{q-1-l}\}$ subset of $V'' \setminus \{j\}$, consider the set $V''' = V'' \setminus \{k_1, \cdots, k_{q-1-l}\}$. Observe $|V'''| = |V| - q + 1$. Thus by assumption $(V''', E_G(V''') \setminus E_p)$ is connected. At the same time V''' has at least two elements, contains j, and yet j is isolated in V''', establishing a contradiction. ■

We now recount a result from [8] that ties connectivity and rigidity to global rigidity.

Theorem 3.3 *A graph $G=(V,E)$ is globally rigid iff it is $(1,2)$-rigid and $(3,1)$-connected.*

We now define redundant global rigidity, which as noted in section 2 is equivalent to redundant localizability given three or more noncollinear anchors.

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

Definition 3.3 *A graph $G=(V,E)$ is (q,p)-globally rigid if it is globally rigid and for all $0 \leqslant l \leqslant q-1$ and $0 \leqslant k \leqslant p-1$, the induced subgraph obtained by removing any l vertices and k edges is globally rigid.*

In view of theorems 3.1-3.3 one has the following result.

Theorem 3.4 *A graph $G=(V,E)$, with $|V| > q+1$ is (q,p)-globally rigid iff it is rigid after removing any $q-1$ vertices and p edges and is connected after removing any $q+1$ vertices and $p-1$ edges.*

Thus redundant global rigidity has two components: Redundant rigidity and redundant connectivity.

Finally, we come to redundant localizability.

Definition 3.4 *A sensor network with a given set of inter-node distance measurements is (q,p) redundantly localizable if it remains localizable after removing any $q-1$ nodes and any $p-1$ edges.*

Since a network is localizable if and only if it is globally rigid and has three or more noncollinear anchors, the previous theorem immediately yields:

Theorem 3.5 *A two-dimensional sensor network with a given set of anchors and inter-node distance measurements and at least $q+2$ nodes is (q,p) redundantly localizable if and only if its associated graph is rigid after removing any $q-1$ vertices and p edges, connected after removing any $q+2$ vertices and $p-1$ edges, and the network retains three or more noncollinear anchors after removing any $q-1$ vertices.*

The rest of the paper is largely concerned with the characterization of the redundant rigidity.

4 Connections between vertex and edge redundancy

In this section we establish relationships between vertex and edge redundancy. The uniform message is that vertex redundancy conditions are stronger than their edge counterparts.

First we present a result concerning redundant rigidity.

Theorem 4.1 *A graph $G=(V,E)$, with $|V| \geqslant p+q+3$ that is $(q+s,p)$-rigid for some $s>0$, is $(q,p+s)$-rigid.*

Proof: Consider any graph $G'=(V,E')$ where E' is a subset of E and $|E'|=|E-p+1|$. By definition G' is $(p+s,1)$-rigid. By Lemma 4 of [13] G' is $(q,s+1)$-rigid. Thus by theorem 3.1 G must be $(q,p+s)$-rigid. ■

Next we address connectivity.

Theorem 4.2 *A (q,p)-connected graph is $(q-s,p+s)$-connected for all $0 \leqslant s < q$.*

Proof: All it takes is to prove the result for $s=1$. Then a simple induction will complete the proof. We first assert that if a graph $G=(V,E)$ is connected then for any E_1 a set of edges incident on the elements of V, $G=(V,E \bigcup E_1)$ is connected. Suppose now G is (q,p)-connected. To prove the result we need to show that the induced subgraph $G'=(V',E')$ obtained by removing any $q-2$ vertices and $p-1$ edges is $(2,2)$ connected. By hypothesis G' is $(2,1)$-connected. To establish a contradiction suppose the removal of an edge, $e \in E'$, makes $G''=(V',E' \setminus e)$ not connected. This must be false as by Bolobas, [3], Theorem 2.5, in the $(2,1)$-connected graph G' any pair of vertices must have at least two disjoint paths connecting them. The result follows. ■

We conclude this section by the following theorem that directly follows from Theorems 4.1 and 4.2, together with the characterization of redundnant global rigidity in terms of redundant rigidity and connectivity of Theorem 3.4.

Theorem 4.3 *A graph $G=(V,E)$, with $|V| \geqslant p+q+3$ that is $(q+s,p)$-globally rigid for some $s>0$, is $(q,p+s)$-globally rigid.*

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

5 Redundant Edge Rigidity: The Laman Approach

This section focuses on providing a *Laman type* necessary and sufficient condition for $(1,p)$-rigidity. First we need a definition of *minimal (1,p)-rigidity.* Unsurprisingly, when $p=1$, it is consistent with the standard notion of (nonredundant) minimal rigidity.

Definition 5.1 *A graph $G=(V,E)$ is minimally $(1,p)$-rigid if it is $(1,p)$-rigid and loses rigidity after the removal of any set of p edges.*

We first provide a necessary and sufficient condition for minimal $(1,p)$-rigidity.

Theorem 5.1 *A graph $G=(V,E)$ is minimally $(1,p)$-rigid if and only if both the conditions below hold:*

(a) $|E|=2|V|-4+p$.
(b) For any $X\subset V$, $1<|X|<|V|$, $i_G(X)\leqslant 2|X|-3$.

Proof: For necessity assume that G is minimally $(1,p)$-rigid. To prove necessity of (a), observe that if $|E|<2|V|-4+p$ then removal of $p-1$ edges will result in a subgraph with fewer than $2|V|-3$ edges rendering the subgraph nonrigid. Now suppose $|E|>2|V|-4+p$, and the graph is rigid. Then the removal of at least one set of p edges will retain rigidity. Hence for minimal $(1,p)$-rigidity (a) is necessary. Suppose now the graph is minimally $(1,p)$-rigid and (a) holds but (b) is violated. This means there is a $X\subset V$, $1<|X|<|V|$, such that $i_G(X)>2|X|-3$. Now $(1,p)$-rigidity must imply that every vertex has at least $p+1$ edges incident on it. Now remove $p-1$ edges from any such vertex not in X. Call the resulting edge set E'. Clearly by (a) $|E'|=2|V|-3$. The graph $G'=(V,E')$ by assumption is minimally rigid. Observe that the edges removed from E to obtain E' cannot be in the subgraph induced by X in G. Thus $i_{G'}(X)=i_G(X)>2|X|-3$, establishing a contradiction by violating Laman's theorem.

To prove sufficiency observe that as (a) is necessay, if G is $(1,p)$-rigid then it is minimally so. Thus to prove sufficiency we need simply show that (a,b) imply $(1,p)$-rigidity. Indeed remove any $p-1$ edges from E. Call the resulting edge set E'. Clearly $|E'|=2|V|-3$. Further because of (b), for the graph $G'=(V,E')$, and all $X\subset V$, $1<|X|<|V|$, one has $i_{G'}(X)\leqslant i_G(X)\leqslant 2|X|-3$. Hence G' is minimally rigid.

■

This result is in fact quite powerful, in that beyond a precise edge count, all it requires is that Laman's condition be satisfied by the graph induced by any proper subset of V.

Rigidity, as opposed to minimal rigidity, is characterized through Laman's theorem using the property that for a rigid graph $G=(V,E)$ there must exist an $E'\subset E$, such that $G=(V,E')$ is minimally rigid. The example depicted in figure 3(a) shows however, that *this is not true in general* for $(1,p)$-rigidity.

To analyze the graph $G=(V,E)$, depicted in figure 3(a) observe that every vertex in the first and last rows of this nine vertex graph has an edge to every vertex in the second row. Thus in all there are 18 edges in this graph. Obviously, as

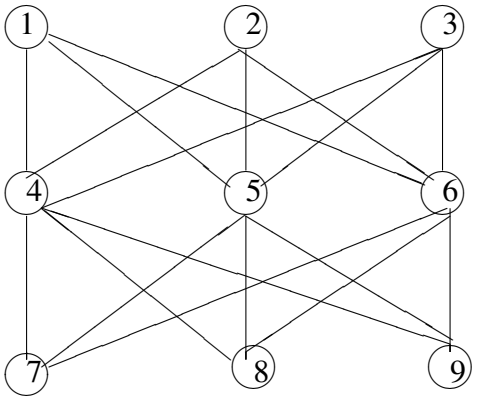
$$18>2\times 9-4+2,$$

this graph cannot be minimally $(1,2)$-rigid. It is also clear that there is no $E'\subset E$ such that (V,E') is minimally $(1,2)$-rigid. This is so as were such an E' to exist it, $E\backslash E'$ must have precisely two elements in it. It is clear that the removal of any edge from G leaves at least one vertex with degree equal to two. Then no matter what further edge is removed, the resulting subgraph cannot be $(1,2)$-rigid.

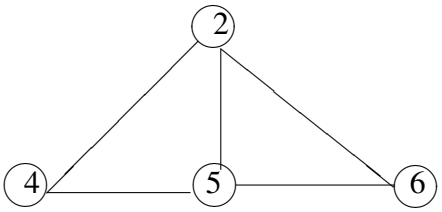
We now assert that G is in fact $(1,2)$-rigid. To see this first observe that the graph induced by the vertices $\{1,\cdots,6\}$ and that induced by $\{4,\cdots,9\}$ are each minimally rigid. Indeed the graph in figure 3(b) is minimally rigid. The graph induced by $\{1,\cdots,6\}$ can be built from this by two edge-splitting operations: First add vertex 1 with its three edges. Remove edge e_{45}. Second add vertex 3 with its three edges. Remove edge e_{65}. The minimal rigidity of the graph induced by $\{4,\cdots,9\}$ similarly follows.

Also note that there is a complete symmetry between the edges, in that if we show that the graph obtained by removing the edge e_{14} is rigid, this implies that a graph obtained by removing any single edge is rigid. Now remove the edge e_{14}. Since the

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST



(a)



(b)

Fig. 3. (a) A $(1,2)$-rigid graph that does not contain a minimally $(1,2)$-rigid subgraph; (b) A graph used in proving claim for graph of (a).

graph induced by $\{4,\cdots,9\}$ is minimally rigid, the graph $(V,E\setminus\{e_{14},e_{25},e_{36}\})$ is minimally rigid, being obtainable by three edge splitting Henneberg operations, [7]. Thus the graph $(V,E\setminus\{e_{14}\})$ is rigid. Thus indeed the graph in figure 3 is $(1,2)$-rigid.

Despite the above observation that a $(1,p)$-rigid graph may not contain a minimally $(1,p)$-rigid graph, the following theorem does hold.

Theorem 5.2 *A graph $G=(V,E)$ is $(1,p)$-rigid if there is an $E'\subset E$, such that $G=(V,E')$ is minimally $(1,p)$-rigid.*

Proof: Suppose there is an $E'\subset E$, such that $G=(V,E')$ is minimally $(1,p)$-rigid. Consider any $E_p\subset E$, comprising $p-1$ elements. Observe $|E_p\cap E'|<p$. Thus $G=(V,E'\setminus E_p)$ is rigid. Consequently as adding edges to a rigid graph retains rigidity and $E'\setminus E_p\subset E\setminus E_p$, it follows that $G=(V,E\setminus E_p)$ is rigid, proving the result. ■

Evidently then there are two distinct classes of $(1,p)$-rigid graphs. The first which can be made minimally $(1,p)$-rigid after removing some edges. The other is exemplified by the example of figure 3(a). It is thus helpful to formalize this distinction through two separate classifications of $(1,p)$-rigidity.

Definition 5.2 *A graph $G=(V,E)$ is strongly $(1,p)$-rigid if there exists an $E'\subset E$, such that $G=(V,E')$ is minimally $(1,p)$-rigid.*

By contrast we have the class represented by the example in figure 3(a).

Definition 5.3 *A graph $G=(V,E)$ is weakly $(1,p)$-rigid if it is $(1,p)$-rigid but for every $E'\subset E$, $G=(V,E')$ is not minimally $(1,p)$-rigid.*

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

Evidently using Theorem 5.2 one arrives at the Laman type characterization of strong $(1,p)$-rigidity provided in Theorem 5.3 below, that is a direct analogy of Theorem 2.1. On the other hand a Laman type characterization of weak $(1,p)$-rigidity remains open.

Theorem 5.3 *A graph $G=(V,E)$ is strongly $(1,p)$-rigid if and only if there is an $E'\subset E$, such that the subgraph $G'=(V,E')$ obeys both the conditions below:*

(a) $|E'|=2|V|-4+p$.
(b) For any $X\subset V$, $1<|X|<|V|$, $i_{G'}(X)\leqslant 2|X|-3$.

6 Redundant Edge Rigidity: The Lovasz-Yemeni Approach

In this section, we characterize $(1,p)$-rigidity using the Lovasz-Yemeni approach. As in section 5 we will obtain the result attending first to minimally $(1,p)$-rigid graphs.

Lemma 6.1 *Let $G=(V,E)$ be a minimally $(1,p)$-rigid graph. Then $r_p(G)=2|V|-4+p$, where $r_p(G)$ is defined in Definition 2.3.*

Proof: As the graph is minimally $(1,p)$-rigid, from Theorem 5.1, $|E|=2|V|-4+p$. Now consider $F_1,\cdots,F_l$,

$$l\geqslant p, \tag{6.2}$$

with each F_i a nonempty subset of E. Suppose,

$$\bigcup_{i=1}^{l} F_i = E. \tag{6.3}$$

Then from Definition 2.1,

$$P=\{G_i=(V_G(F_i),F_i)\}_{i=1}^{l} \tag{6.4}$$

is an l-AD of G. Given that $l\geqslant p$, any l-AD is also a p-AD. Thus (6.4) is a p-AD under (6.2-6.4). Henceforth assume (6.2-6.4) is in force. We will consider two cases that cover all possibilities.

Case I: For some i

$$V_G(F_i)=V.$$

Observe that for any $j\neq i$, and nonempty F_j, $2V_G(F_j)-3\geqslant 1$. Thus as (6.2) holds the corresponding $r(P)\geqslant(2|V|-3)+l-1\geqslant 2|V|-4+p$.

Case II: There holds:

$$V_G(F_i)\neq V\forall i \tag{6.5}$$

Call the class of $F=\{F_1,\cdots,F_l\}$ that lead to (6.2-6.5), $\mathscr{F}$. Observe from (6.5) that for each i, and $F\in\mathscr{F}$, $|V_G(F_i)|<|V|$. By identifying $V_G(F_i)$ with X in Theorem 5.1, we see that $|F_i|\leqslant 2|V_G(F_i)|-3$, i.e. each F_i is a maximally independent subset in G_i. Consequently, by Lemma 2.2, for each i, $G_i=(V_G(F_i),F_i)$ obeys $r_1(G_i)=|F_i|$. Then there holds:

$$r_p(G)=\min_{F\in\mathscr{F}}\left\{\sum_{i=1}^{l} r_1((V_G(F_i),F_i))\right\}=\min_{F\in\mathscr{F}}\left\{\sum_{i=1}^{l}|F_i|\right\}\geqslant|E|=2|V|-4+p. \tag{6.6}$$

It remains to show that there is a p-AD whose index is exactly $2|V|-4+p$. Indeed label the elements of E as $e_1,\cdots,e_{2|V|-4+p}$, and choose $P''=\{(V_G(\{e_i\}),\{e_i\})\}_{i=1}^{2|V|-4+p}$. Then as for each i, $2|V_G(\{e_i\})|-3=1$, $r(P'')=|E|=2|V|-4+p$. ∎

We next derive a result that shows that $r_p(G)\geqslant 2|V|-4+p$ is a sufficient condition for $(1,p)$-rigidity.

Lemma 6.2 *Let $G=(V,E)$ be a graph with $r_p(G)\geqslant 2|V|-4+p$. Then G is $(1,p)$-rigid.*

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

Proof: Suppose the graph is not $(1,p)$-rigid. Then there exists an E' obtained by removing an edge set E_p comprising $p-1$ edges, from E such that $G'=(V,E')$ is nonrigid. Thus, from Theorem 2.4, there is an AD of G', call it P, such that $r(P) \leqslant 2|V|-4$. Now consider

$$P' = \left(\bigcup_{e \in E_p} \{(V_G(\{e\}),\{e\})\} \right) \bigcup P.$$

Observe $|V_G(\{e\})| = 2$ for each $e \in E_p$. As P' is a p-AD of G, from (2.1), we obtain:

$$r_p(G) \leqslant r(P') = r(P) + \sum_{e \in E_p} (2|V_G(\{e\})| - 3) \leqslant r(P) + p - 1 \leqslant 2|V| - 5 + p$$

establishing a contradiction.

■

Combining these two Lemmas we obtain the following necessary and sufficient condition for minimal $(1,p)$-rigidity.

Theorem 6.1 *A graph $G=(V,E)$ is minimally $(1,p)$-rigid iff $r_p(G) \geqslant |E| = 2|V|-4+p$.*

Proof: By Theorem 5.1, condition (a) and Lemma 6.1, minimal $(1,p)$-rigidity implies $r_p(G) = |E| = 2|V|-4+p$. Conversely, suppose that $r_p(G) \geqslant |E| = 2|V|-4+p$. By Lemma 6.2, G is $(1,p)$-rigid. Now argue as in the proof of Theorem 5.1: as $|E| = 2|V|-4+p$, the removal of any set of p edges will give a graph with $2|V|-4$ vertices, i.e. a nonrigid graph. This establishes minimal rigidity. ■

The question remains whether $r_p(G) = 2|V|+p-4$ is a necessary condition for *nonminimal* $(1,p)$-rigidity (assuming that $|E| \neq 2|V|+p-4$). The graph $G=(V,E)$ in figure 4 serves as a counter example.

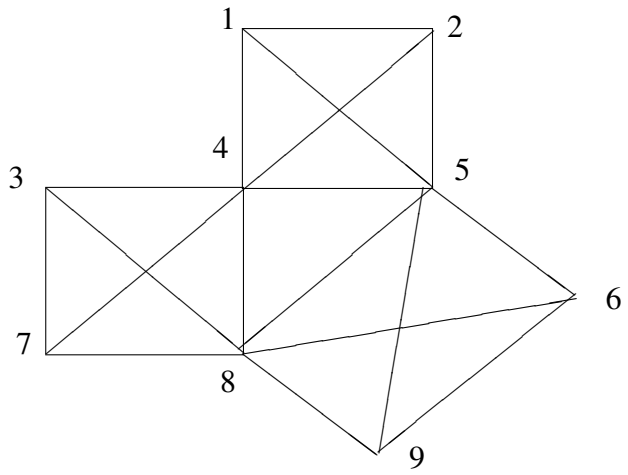


Fig. 4. A $(1,2)$-rigid graph with $r_2(P) < 2|V|-2$.

Consider the three edge partitions: $E_1 = E_G(\{1,2,4,5\})$, $E_2 = E_G(\{3,7,4,8\})$ and $E_3 = E_G(\{8,9,6,5\})$. $P = \{(V_G(E_i),E_i)\}_{1=1}^3$ is a 2-AD for G. In this case:

$$r(P) = 3 \times (8-3) = 15 < 2 \times 9 - 2 = 16.$$

The graph is clearly rigid after removing any single edge other than from the set e_{45}, e_{48}, e_{58}. Now without sacrificing generality remove the edge e_{45}. Observe it can be reconstructed by an edge-splitting operation on $(V \setminus \{2\}, E_G(V \setminus \{2\}))$. Thus this graph is $(1,2)$-rigid.

Nonetheless the obvious Lovasz-Yemeni type necessary condition does hold for *strong* $(1,p)$-rigidity, as we now see.

Theorem 6.2 *A graph $G=(V,E)$ with $r_p(G) \geqslant 2|V|-4+p$ is $(1,p)$-rigid. It is strongly $(1,p)$-rigid only if $r_p(G) \geqslant 2|V|-4+p$..*

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

Proof: The first statement follows from Lemma 6.2.

We will use induction to prove the second claim. Clearly from Theorem 2.4 the graph $G=(V,E)$ is $(1,1)$-rigid only if $r_1(G) \geqslant 2|V|-4+1$. Now suppose for all $1 \leqslant l \leqslant p-1$ the graph $G=(V,E)$ is strongly $(1,l)$-rigid only if $r_l(G) \geqslant 2|V|-4+l$.

Suppose the graph $G=(V,E)$ is $(1,p)$-rigid but $r_p(G) < 2|V|-4+p$. Then from Lemma 6.1, it is not minimally rigid. From Theorem 5.2 there is an $E' \subset E$, such that $G'=(V,E')$ is minimally $(1,p)$-rigid. Call $F=E \setminus E'$. Now consider any p-AD, $P=\{(V_G(E_1),E_1)\cdots,(V_G(E_k),E_k)\}$ of G. By definition $k \geqslant p$. Further:

$$\bigcup_{i=1}^{k} E_i = E.$$

Thus:

$$\bigcup_{i=1}^{k} (E_i \setminus F) = E'.$$

Without sacrificing generality label the nonempty sets among $E_i \setminus F$ as $E_1 \setminus F, \cdots E_m \setminus F$, with $1 \leqslant m \leqslant k$.

Then $P'=\{(V_G(E_1 \setminus F),E_1 \setminus F)\cdots,(V_G(E_m \setminus F),E_m \setminus F)\}$ is an m-AD of G'. Observe also that:

$$|V_G(E_i \setminus F)| \leqslant |V_G(E_i)| \ \ \forall 1 \leqslant i \leqslant m \tag{6.7}$$

and

$$|V_G(E_i)| \geqslant 2 \ \ \forall m+1 \leqslant i \leqslant k. \tag{6.8}$$

thus one obtains

$$\begin{aligned} r(P) &= \sum_{i=1}^{k} [2|V_G(E_i)|-3] \\ &\geqslant \sum_{i=1}^{m} [2|V_G(E_i \setminus F)|-3] + \sum_{i=m}^{k} [2|V_G(E_i)|-3] \end{aligned}$$

or

$$r(P) \geqslant r(P') + k - m. \tag{6.9}$$

Now supposoe $m < p$ (the contrary case is considered below). Then the graph $G(V,E')$ is $(1,m)$-rigid by the inductive hypothesis:

$$r(P') \geqslant 2|V|-4+m.$$

Thus from (6.9) there holds,

$$r(P) \geqslant r(P') + k - m \geqslant 2|V|-4+m+k-m = 2|V|-4+k \geqslant 2|V|-4+p.$$

On the other hand if $m \geqslant p$ then $r(P)$ is still no less than $r(P')$ as (6.7) and (6.8) continue to hold. Consequently from Lemma 6.1,

$$r(P) \geqslant r(P') \geqslant 2|V|-4+p. \tag{6.10}$$

The result follows.

■

Thus while both Lovasz-Yemeni and Laman type necessary and sufficient conditions exist for minimal $(1,p)$-rigidity, and are necessary for strong $(1,p)$-rigidity, these only provide sufficient conditions for weak $(1,p)$-rigidity.

Recall however, that the direct subject of investigation in this paper is redundant global rigidity, which beyond redundant rigidity also require redundant connectivity. It is instructive to note that the example of Fig. 3 is *(3,1)-connected, though that of figure 4 is not*. It is also possible to show that the graph in Fig. 3 does satisfy the corresponding Lovasz-Yemeni sufficient condition for $(1,2)$-rigidity. Indeed the next section demonstrates that a $(1,p)$-rigid graph that is not necessarily strongly $(1,p)$-rigid but has an additional redundant connectivity property must have $r_p(P) \geqslant 2|V|-4+p$.

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

7 Lovasz-Yemeni type condition for redundant rigidity with redundant connectivity

This section considers levels of connectivity needed for a $(1,p)$-rigid graph to satisfy $r_p(P) \geqslant 2|V| - 4 + p$. At the end of this section we will tie the results here to redundant global rigidity.

We first present a Lemma that shows that for any $p \geqslant 1$, at least one p-minimizing AD of $G = (V,E)$ must have underlying edge sets that partition, rather than merely cover, E. This theorem thus extends a simplifying conclusion of Theorem 2.4 to the case of $p > 1$.

Lemma 7.1 *Suppose $P = \{(V_G(E_i), E_i)\}_{i=1}^k$, a k-AD of $G = (V,E)$. Then there are $\bar{E}_1, \cdots, \bar{E}_k$, that partition E, such that for $\bar{P} = \{(V_G(\bar{E}_i), \bar{E}_i)\}_{i=1}^k$, there holds:*

$$r(\bar{P}) \leqslant r(P).$$

Proof: Since P is a k-AD, the E_i are all nonempty and form a cover of E. Thus there exist mutually disjoint, nonempty

$$\bar{E}_i \subset E_i$$

such that the $\bar{E}_i$ partition E. Then the result follows by noting that $V_G(\bar{E}_i) \subset V_G(E_i)$.

■

Because of this we will assume henceforth that the edge covers underlying the ADs we consider are disjoint. Since our goal here is to tie Lovasz-Yemeni type conditions to a coupling of redundant connectivity and redundant rigidity, the next lemma records a key consequence of redundant connectivity.

Lemma 7.2 *Suppose $G = (V,E)$ is $(q,1)$-connected. Consider $P = \{(V_G(E_i), E_i)\}_{i=1}^k$, a k-AD of G, with E_i mutually disjoint and $k \geqslant 2$. Suppose for some i that $|V_G(E_i)| \geqslant q$ and $|V_G(E \setminus E_i)| \geqslant q$. Then $|V_G(E_i) \cap V_G(E \setminus E_i)| \geqslant q$.*

Proof: Consider Fig. 5, which depicts the three vertex sets, $V_G(E_i) \setminus V_G(E \setminus E_i)$, $V_G(E_i) \cap V_G(E \setminus E_i)$ and $V_G(E \setminus E_i) \setminus V_G(E_i)$. Observe all paths from vertices in the leftmost set to those in the rightmost are through the middle set. To establish a contradiction suppose the middle set has fewer than q vertices, then the removal of these vertices makes the middle set empty, making the resulting graph not connected. This contradicts the fact that the original graph is $(q,1)$-connected.

■

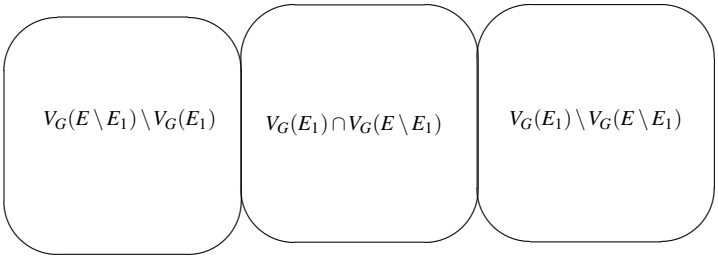


Fig. 5. Illustration of the proof of Lemma 7.2.

Next we record a simple lemma that will be exploited in the sequel.

Lemma 7.3 *Consider a set X and subsets $X_1, \cdots, X_k$, $k > 1$ that cover X. Suppose for all i,*

$$X_i \subset \bigcup_{j \neq i} X_j. \tag{7.11}$$

Then there holds:

$$2|X| \leqslant \sum_{i=1}^{k} |X_i| \tag{7.12}$$

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

Proof:

Since $X_1, \cdots, X_k$ cover X, (7.11) ensures that every element of X appears in at least two sets among the X_i. ■

Much of the analysis in the remainder of this section hinges on the following pivotal Lemma.

Lemma 7.4 *Consider a set V and subsets $V_1, \cdots, V_k$, $k > 1$ that cover V. Suppose for some q and all i, $\left|V_i \cap \{\bigcup_{j \neq i} V_j\}\right| \geqslant q$. Then there holds:*

$$2\sum_{i=1}^{k} |V_i| - 2|V| \geqslant qk. \tag{7.13}$$

Proof: Define

$$X_i = V_i \cap \left\{ \bigcup_{j \neq i} V_j \right\} \tag{7.14}$$

and

$$S_i = V_i \setminus X_i. \tag{7.15}$$

Also denote:

$$S = \bigcup_{i=1}^{k} S_i, \tag{7.16}$$

and

$$X = \bigcup_{i=1}^{k} X_i, \tag{7.17}$$

From (7.14) and the theorem statement it is evident that

$$|X_i| \geqslant q. \tag{7.18}$$

From (7.14) and (7.15) it is also clear that for all i, j,

$$S_i \cap X_j = \{\phi\}, \tag{7.19}$$

for all $i \neq j$,

$$S_i \cap S_j = \{\phi\}, \tag{7.20}$$

and for all i,

$$|V_i| = |S_i| + |X_i|. \tag{7.21}$$

From (7.19)-(7.21) one can verify after some manipulation that for all i

$$V_i \cap \left\{ \bigcup_{j \neq i} V_j \right\} = X_i \cap \left\{ \bigcup_{j \neq i} X_j \right\}.$$

Consequently X_i obey (7.11) and Lemma 7.3 applies.

Further, the S_i partition S, while the X_i cover X, and $V = S \cup X$. Then because of (7.19) and (7.20) there holds:

$$\begin{aligned} |V| &= \left| \left\{ \bigcup_{i=1}^{k} S_i \right\} \bigcup \left\{ \bigcup_{i=1}^{k} X_i \right\} \right| \\ &= \sum_{i=1}^{k} |S_i| + |X|. \end{aligned}$$

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

Thus, because of (7.12) and (7.19), and Lemma 7.3, there holds:

$$\begin{aligned}\sum_{i=1}^{k}|V_i|-|V| &= \sum_{i=1}^{k}|V_i|-\sum_{i=1}^{k}|S_i|-|X| \\ &= \sum_{i=1}^{k}|S_i|+\sum_{i=1}^{k}|X_i|-\sum_{i=1}^{k}|S_i|-|X| \\ &= \sum_{i=1}^{k}|X_i|-|X| \\ &\geqslant \frac{1}{2}\sum_{i=1}^{k}|X_i|.\end{aligned}$$

Thus, because of (7.18) we obtain:

$$2\sum_{i=1}^{k}|V_i|-2|V| \geqslant \sum_{i=1}^{k}|X_i| \geqslant qk.$$

■

We then have the first result connecting Yemeni-Lovasz type conditions to $(1,p)$-global rigidity for $p \leqslant 4$.

Theorem 7.1 *Suppose, for integer p with* $2 \leqslant p \leqslant 4$, $G=(V,E)$ *is* $(3,p-1)$*-connected and* $(1,p)$*-rigid. Then* $r_p(G) \geqslant 2|V|-4+p$.

Proof: We will use induction. To obtain the basis step consider first $p=2$. Suppose $G=(V,E)$ is $(3,1)$-connected and $(1,2)$-rigid. (In fact, this means that G is globally rigid; we will explore this issue later in more detail). Suppose for $k \geqslant 2$, $P=\{(V_G(E_i),E_i)\}_{i=1}^{k}$ is a k-AD of G, with E_i partitioning E. Then by Lemma 7.1 $r_2(G)$ is the index of one such k-AD.

Suppose first that for some i, $|E_i|=1$. Without loss of generality assume $i=1$. Since G is $(1,2)$-rigid, $G'=(V,E\setminus E_1)$ is rigid. Since the E_i are disjoint, for all $i>1$ $E_i \setminus E_1 = E_i$ are therefore all nonempty and partition $E\setminus E_1$. Thus, $P_1=\{(V_G(E_i\setminus E_1),E_i\setminus E_1)\}_{i=2}^{k}$, is a $(k-1)$-AD of G'. Since G' is rigid, by Theorem 2.4

$$r(P') \geqslant 2|V|-3.$$

Hence:

$$r(P)=r(P')+2|V_G(E_1)|-3 \geqslant 2|V|-2,$$

i.e all such P have indices that exceed or equal the hypothesized value of $r_2(G)$.

There remains the case where all $|E_i|>1$. In this case for all i, $|V_G(E_i)| \geqslant 3$. Further, as the E_i partition E, for all i there holds

$$V_G(E\setminus E_i)=V_G\left(\bigcup_{j\neq i}E_j\right)=\bigcup_{j\neq i}V_G(E_j). \tag{7.22}$$

As the graph is $(3,1)$-connected, by Lemma 7.2, for all i,

$$|V_G(E\setminus E_i)\cap V_G(E_i)| \geqslant 3. \tag{7.23}$$

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

Thus, identifying $V_i = V_G(E_i)$, conditions of Lemma 7.4 hold with $q = 3$. Consequently,

$$\begin{aligned} r(P) &= \sum_{i=1}^{k} (2|V_i| - 3) \\ &= 2\sum_{i=1}^{k} |V_i| - 3k \\ &\geqslant 2|V| + 3k - 3k \\ &> 2|V| - 2 \end{aligned}$$

Now for the induction step, suppose the result holds for some p, $2 \leqslant p < 4$. Suppose $G = (V,E)$ is $(3,p)$-connected and $(1,p+1)$-rigid. Consider for $k \geqslant p+1$, $P = \{(V_G(E_i),E_i)\}_{i=1}^{k}$, a k-AD of G, with E_i partitioning E. Then by Lemma 7.1 $r_{p+1}(G)$ is the index of one such k-AD.

Suppose first that for some i, $|E_i| = 1$. Without loss of generality assume $i = 1$. Since G is $(3,p)$-connected and $(1,p+1)$-rigid, $G' = (V, E \setminus E_1)$ is $(3,p-1)$-connected and $(1,p)$-rigid. Since the E_i are disjoint, for all $i > 1$ $E_i \setminus E_1 = E_i$ are therefore all nonempty and partition $E \setminus E_1$. Thus, $P_1 = \{(V_G(E_i \setminus E_1), E_i \setminus E_1)\}_{i=2}^{k}$, is a $(k-1)$-AD of G'. By the induction hypothesis:

$$r(P') \geqslant 2|V| - 4 + p.$$

Hence:

$$r(P) = r(P') + 2|V_G(E_1)| - 3 \geqslant 2|V| - 4 + p + 1,$$

i.e all such P have indices that exceed or equal the hypothesized value of $r_{p+1}(G)$. There remains the case where all $|E_i| > 1$. In this case for all i, $|V_G(E_i)| \geqslant 3$. Further, as the E_i partition E, for all i, (7.22) holds. Further, a $(3,p)$-connected graph is also $(3,1)$-connected. Then, by Lemma 7.2, for all i, (7.23) holds.

Thus, identifying $V_i = V_G(E_i)$, conditions of Lemma 7.4 hold with $q = 3$. Consequently, as $p < 4$,

$$\begin{aligned} r(P) &= \sum_{i=1}^{k} (2|V_i| - 3) \\ &= 2\sum_{i=1}^{k} |V_i| - 3k \\ &\geqslant 2|V| + 3k - 3k \\ &\geqslant 2|V| + 1 + p - 4 \end{aligned}$$

Thus the result follows. ∎

There remains the question whether the result of Theorem 7.1 will hold for $p > 4$. We now present a counterexample for $p = 5$.

Example 7.1 *The graph in question, $G = (V,E)$ has $|V| = 33$, and has a vertex cover $V_1, \cdots, V_6$, such that each $|V_i| = 7$, each subgraph $(V_i, E_G(V_i))$ is a K_7 graph, and the sets $E_i = E_G(V_i)$ consitute a partition of E. Further, defining*

$$X_i = V_i \cap \left(\cup_{j \neq i} V_j\right), \tag{7.24}$$

one has $X_1 = \{1,2,3\}$, $X_2 = \{1,4,5\}$, $X_3 = \{2,6,7\}$, $X_4 = \{3,8,9\}$, $X_5 = \{4,6,8\}$ and $X_6 = \{5,7,9\}$. In other words V_i comprises X_i together with four other vertices that are not in any other V_i. Thus there are 24 vertices each belonging to just one V_i, and 9 which are in exactly two.

To see that this choice of X_i is consistent with the fact that the E_i partition E, observe for all $i \neq j$,

$$\left|X_i \cap X_j\right| \leqslant 1. \tag{7.25}$$

Further vertices in each V_i connect to the others through the elements of X_i. Thus, indeed this defintion is consistent with the requirement that the E_i are disjoint.

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

In effect then G comprises six K_7 subgraphs, connected through the vertices in the X_i. Each V_i has three vertices through which it connects to other V_j, there being one vertex in common with each of three different V_j.

It is shown in the appendix that G is both $(3,4)$-connected and $(1,5)$-rigid. Since each subgraph $(V_i, E_G(V_i))$ is K_7, it also follows that $V_i = V_G(E_i)$. Thus as the E_i partition E, we have that

$$P = \{(V_i, E_G(V_i))\}_{i=1}^{6}$$

is a 6-AD of G. Hence,

$$r_5(G) \leqslant r(P) = 6(14-3) = 66 = 2|V| < 2|V| - 4 + 5,$$

and Theorem 7.1 is violated for $p = 5$.

Nonetheless we now show that for $p > 4$, a stronger redundant connectivity condition suffices for $r_p(G) \geqslant 2|V| - 4 + p$ to hold given $(1,p)$-rigidity.

Theorem 7.2 *Suppose, for integer $p \geqslant 3$, $G = (V,E)$ is $(4, p-2)$-connected and $(1,p)$-rigid. Then $r_p(G) \geqslant 2|V| - 4 + p$.*

Proof: From Theorem 4.2, a $(4, p-2)$-connected graph is also $(3, p-1)$-connected. Thus from Theorem 7.1, the result holds for $3 \leqslant p \leqslant 4$.

To use induction suppose the result holds for all $3 \leqslant p \leqslant n$, $n \geqslant 4$. Suppose $G = (V,E)$ is $(4, n-1)$-connected and $(1, n+1)$-rigid. Consider for $k \geqslant n+1$, $P = \{(V_G(E_i), E_i)\}_{i=1}^{k}$, a k-AD of G, with E_i partitioning E. Then by Lemma 7.1 $r_{n+1}(G)$ is the index of one such k-AD.

Suppose first that for some i, $|V_G(E_i)| = 3$. Then $|E_i| \leqslant 3$. Without loss of generality assume $i = 1$.

Consider $G' = (V, E \setminus E_1)$. Since the E_i are disjoint, for all $i > 1$ $E_i \setminus E_1 = E_i$. Thus, $P_1 = \{(V_G(E_i \setminus E_1), E_i \setminus E_1)\}_{i=2}^{k}$, is a $(k-1)$-AD of G'.

First suppose $n = 4$. Since by the induction hypothesis G is $(1,5)$-rigid, G' is $(1,2)$-rigid. Further as G is $(4,3)$-connected, by Theorem 7.1 it is $(3,4)$-connected. Consequently, G' is $(3,1)$-connected.

Thus by Theorem 7.1, applying the case of $p = 2 = n-2$,

$$r(P') \geqslant 2|V| - 4 + n - 2. \tag{7.26}$$

Hence as $|V_G(E_1)| = 3$,

$$r(P) = r(P') + 2|V_G(E_1)| - 3 \geqslant 2|V| - 6 + n + 3 = 2|V| - 4 + (n+1), \tag{7.27}$$

i.e all such P have indices that exceed or equal the hypothesized value of $r_{n+1}(G)$.

Now consider $n > 4$, still with $|V_G(E_1)| = 3$. In this case as G is $(4, n-1)$-connected and $(1, n+1)$-rigid, G' is $(4, n-4)$-connected and $(1, n-2)$-rigid. As $n - 4 \geqslant 1$, by the induction hypothesis (7.26) still holds. Hence as $|V_G(E_i)| = 3$, (7.27) also holds and the requirement of the theorem is met.

Thus when there is a $|V_G(E_i)| = 3$, the requirements of the theorem are met. An identical argument given in the proof of Theorem 7.1 also dispenses with the case when for some i, $|V_G(E_i)| = 2$ as this is tantamount to $|E_i| = 1$. There remains the case where for all i, $|V_G(E_i)| \geqslant 4$. As in Theorem 7.1 (7.22) holds. Further, as $n \geqslant 4$, G is $(4,1)$-connected. Then by Lemma 7.2, for all i,

$$|V_G(E \setminus E_i) \cap V_G(E_i)| \geqslant 4. \tag{7.28}$$

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

Thus, identifying $V_i = V_G(E_i)$, conditions of Lemma 7.4 hold with $q = 4$. Consequently, as $k \geqslant n+1$,

$$\begin{aligned} r(P) &= \sum_{i=1}^{k} (2|V_i| - 3) \\ &= 2\sum_{i=1}^{k} |V_i| - 3k \\ &\geqslant 2|V| + 4k - 3k \\ &= 2|V| + k \\ &\geqslant 2|V| + n - 3. \end{aligned}$$

The result follows. ∎

As global rigidity is equivalent to the combined properties of $(3,1)$-connectivity and $(1,2)$-rigidity, in view of the two theorems in this section we have the following key result.

Theorem 7.3 *A graph $G = (V,E)$ is $(1,p)$-globally rigid if it is $(3,p)$-connected and $r_{p+1}(G) \geqslant 2|V| - 3 + p$. For $1 \leqslant p \leqslant 3$ it is $(1,p)$-globally rigid only if $r_{p+1}(G) \geqslant 2|V| - 3 + p$. For $p \geqslant 4$ a $(4, p-2)$-connected $G = (V,E)$ is $(1,p)$-globally rigid only if $r_{p+1}(G) \geqslant 2|V| - 3 + p$.*

Proof: Observe first that $(1,p)$-global rigidity is equivalent to global rigidity being preserved after the deletion of $p-1$ edges, or having $(3,1)$-connectivity and $(1,2)$-rigidity after the deletion of $p-1$ edges, or having initially $(3,p)$-connectivity and $(1,p+1)$-rigidity. By Theorem 6.2, $r_{p+1}(G) \geqslant 2|V| - 4 + (p+1)$ ensures $(1,p+1)$-rigidity and so the first claim of the theorem holds. The second and third claim follow immediately from Theorems 7.1 and 7.2. ∎

Given the relation between global rigidity, anchor count and localizability, there is an obvious extension of this to a corresponding characterization of redundantly localizable two-dimensional sensor networks.

We comment that in a sensor network in which loss of nodes is contemplated, it may be that loss of anchor nodes can be ruled out on reliability or other grounds. What is needed for redundant localizability in addition to redundant global rigidity is that after loss of up to q-1 nodes, at least three noncollinear anchor nodes remain. The issue of which nodes fail does not enter the picture when the redundancy is all with respect to edge loss of course.

8 Conclusions and Future Work

This work presents fundamental results on edge-redundant global rigidity, which is essential to the problem of seeking a solution to guaranteing network localizability in the event of link losses. The particular notion that is studied in detail is minimal $(1,p)$-rigidity; for which two different types of characterizations were obtained: Laman type and Lovasz-Yemeni type. For the latter, further connection was established between redundant edge rigidity and redundant connectivity.

As discussed in the beginning of this paper, it is of great importance to also study redundant vertex rigidity, i.e. $(q,1)$-rigidity. Some discussion is provided in the text but its characterization remains largely open. Also, one will need computationally efficient algorithms to check for satisfaction of these conditions for redundant localizability of a network. Moreover, operations are required for one to construct, augment, merge and split such networks, while ensuring the level of redundant localizability remain unchanged. Results along the lines of [1] providing simple sufficient conditions for redundant global rigidity, largely in terms of local properties of graphs, would also be welcome. At some point too, random geometric graphs should be investigated.

References

[1] B. D. O. Anderson, P. N. Belhumeur, T. Eren, D. K. Goldenberg, A. S. Morse, W. Whiteley, and Y. R. Yang. Graphical properties of easily localizable sensor networks. *Wireless Networks*, 15(2):177–191, February 2009.

[2] J. Aspnes, T. Eren, D. Goldenberg, A. S. Morse, W. Whiteley, Y. Yang, B. D. O. Anderson, and P. Belhumeur. A theory of network localization. *IEEE Transactions on Mobile Computing*, 5(12):1663–1678, December 2006.

[3] B. Bolobas. *Modern Graph Theory*. Springer, 1998.

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

[4] R. Connelly. Generic global rigidity. *Discrete and Computational Geometry*, 33:549–563, 2005.

[5] T. Eren, B.D.O. Anderson, A.S. Morse, W. Whiteley, and P.N. Belhumeur. Operations on rigid formations of autonoumous agents. *Communications in Information and Systems*, 3(4):223–258, September 2004.

[6] J. E. Graver, B. Servatius, and H. Servatius. *Combinatorial Rigidity*. Graduate Studies in Mathematics. American Mathematical Society, 1993.

[7] L. Henneberg. In *Die graphische Statik der starren Systeme.*, Leipzig, 1911.

[8] B. Jackson and T. Jordan. Connected rigidity matroids and unique realizations of graphs. *Journal of Combinatorial Theory Series B*, 94:1–29, 2005.

[9] Donald J. Jacobs and Bruce Hendrickson. An algorithm for two-dimensional rigidity percolation: the pebble game. *J. Comput. Phys.*, 137(2):346–365, 1997.

[10] G. Laman. On graphs and rigidity of plane skeletal structures. *Journal of Engineering Mathematics*, 4(4):331–340, October 1970.

[11] L. Lovász and Y. Yemini. On generic rigidity in the plane. *SIAM Journal on Algebraic and Discrete Methods*, 3(1):91–98, 1982.

[12] W. Whiteley. Some matroids from discrete applied geometry. In J. E. Bonin, J. G. Oxley, and B. Servatius, editors, *Matroid Theory*, volume 197 of *Contemporary Mathematics*, pages 171–311. American Mathemtical Society, 1996.

[13] C. Yu and B.D.O. Anderson. Development of redundant rigidity theory for formation control. *International Journal of Robust and Nonlinear Control*, 19(13):1427–1446, 2009.

[14] Y. Liu Z. Yang and X. Li. Beyond trilateration: On the localizability of wireless ad-hoc networks. In *Proc. of the IEEE INFOCOM*, Rio de Janeiro, Brazil, 2009.

Preprint submitted to Automatica
Received February 29, 2012 20:41:05 PST

2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

On the Performance Limit of Sensor Localization

Baoqi Huang, Tao Li, Brian D.O. Anderson, Changbin Yu

Abstract— **In this paper, we analyze the performance limit of sensor localization from a novel perspective. We consider distance-based single-hop sensor localization with noisy distance measurements by Received Signal Strength (RSS). Differently from the existing studies, the anchors are assumed to be randomly deployed, with the result that the trace of the associated Cramér-Rao Lower Bound (CRLB) matrix becomes a random variable. We adopt this random variable as a scalar metric for the performance limit and then focus on its statistical attributes. By the Central Limit Theorems for U-statistics, we show that as the number of anchors goes to infinity, this scalar metric is asymptotically normal. In addition, we provide the quantitative relationship among the mean, the standard deviation, the number of anchors, parameters of communication channels and the distribution of the anchors. Extensive simulations are carried out to confirm the theoretical results. On the one hand, our study reveals some fundamental features of sensor localization; on the other hand, the conclusions we draw can in turn guide us in the design of wireless sensor networks.**

I. Introduction

Location information plays a vital role in the applications of sensor networks, for it is useful to report the geographic origin of events, to assist in target tracking, to achieve geographic aware routing, to manage sensor networks, to evaluate their coverage, and so on. A sensor network generally consists of two kinds of nodes: anchors and sensors. Anchor positions are known *a priori* (e.g., through GPS or manual configurations), while sensor positions are unknown and need to be determined through certain procedures of localization. Up to now, considerable efforts have been invested in developing sensor localization algorithms.

Apart from designing sensor localization algorithms, the analysis of localization performance also gains much attention. Performance studies specific to sensor localization algorithms are realized to evaluate and compare different sensor localization algorithms. More importantly, the performance limit of sensor localization, namely the lower bound for location estimate errors produced by all localization algorithms, provides a theoretically optimal performance no matter what sensor localization algorithm is applied, and thus reflects fundamental impacts of various factors on sensor localization in an algorithm-independent manner. Due to the essence of Cramér-Rao Lower Bound (CRLB), it has been widely used to characterize the performance limit of sensor localization [1].

Most of the existing CRLB analysis is based on given sensor-anchor geometries. In this paper, we analyze the performance limit of single-hop sensor localization from a novel perspective. As commonly used in the literature, we adopt the trace of the associated CRLB matrix as a scalar metric for the performance limit of sensor localization [1]. However, differently from existing CRLB studies which require exact sensor-anchor geometries to compute the deterministic CRLB, we assume that a fixed number of sensors and anchors are randomly deployed in a two-dimensional plane with distance measurements from Received Signal Strength (RSS). Consequently, the trace of the associated CRLB matrix becomes a random variable with respect to the sensor-anchor geometries, and we focus on the statistical attributes of the trace of the CRLB.

The motivations of our study are as follows. In a mobile environment, such as ad-hoc networks, target tracking, Simultaneous Localization and Mapping (SLAM) [2], mobile anchors assisting in sensor localization [3] and so on, it is trivial to concentrate on the localization performance in one particular time instant, whereas it is attractive to grasp the average localization performance over a period of time and in a wide region. Hopefully, this can be solved by our statistically modeling method. Furthermore, the advantages of our study include: (i) it provides some knowledge about how the scalar metric, equivalently the minimal mean square estimation error (MSE), is distributed over all possible sensor-anchor geometries; (ii) the mean of the scalar metric reveals how the average minimal MSE with respect to all possible sensor-anchor geometries evolves with the number of anchors, the parameters of communication channels and the measurement noises; (iii) the ratio of the standard deviation to the mean indicates the sensitivity of the minimal MSE to sensor-anchor geometries; (iv) it not only provides insights into single-hop sensor localization including source localization and target tracking as specific cases, but also as a prototype paves the way for dealing with more complicated scenarios of sensor localization. In summary, statistical sensor-anchor geometry modeling is a powerful method for investigating the performance limit of sensor localization, which is an essential problem for sensor networks. To the best of our knowledge, this method has never been considered.

Essentially, this scalar metric is a function of U-statistics.

B. Huang is with the Australian National University and NICTA Ltd., Canberra, A.C.T., Australia, and also with Shandong Computer Science Center, Jinan, China. Email:Baoqi.Huang@anu.edu.au

B.D.O. Anderson is with the Australian National University and NICTA Ltd., Canberra, A.C.T., Australia. Email:Brian.Anderson@anu.edu.au

T. Li is with the Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China Email:litao@amss.ac.cn

C. Yu is with the Australian National University and NICTA Ltd., Canberra, A.C.T., Australia, and also with Shandong Computer Science Center, Jinan, China. Email:Brad.Yu@anu.edu.au. Correspondence: RSISE Building 115, The Australian National University, Canberra ACT 0200, Australia, Tel: +61 2 61258670, Fax: +61 2 61258660.

978-1-61284-799-3/11/$26.00 ©2011 IEEE

In statistical theory, U-statistics introduced by the seminal paper [4] are a class of important statistics, and are of great significance in estimation theory in that asymptotic properties of both estimators and test statistics have been derived by using the Central Limit Theorems for U-statistics. Based on the theory of U-statistics, we show that as the number of anchors goes to infinity, this scalar metric is asymptotically normal. We provide the quantitative relationship among the mean, the standard deviation, the number of anchors, parameters of communication channels and the distribution of the anchors. Since our results are based on an asymptotic analysis, the conditions under which our results approximate the real situations well are identified.

The remainder of this paper is organized as follows. The next section introduces the problem formulation. Section III presents the main results about statistical attributes of performance limits. Finally, we conclude this paper and shed light on future work in Section V.

II. Problem Formulation

In this section, we formulate the scalar metric of the performance of single-hop sensor localization using RSS measurements and define a random sensor-anchor geometry model. Throughout this paper, we shall use the following mathematical notations: $(\cdot)^T$ denotes transpose of a matrix or a vector; $Tr(\cdot)$ denotes the trace of a square matrix; $Pr\{\cdot\}$ denotes the probability of an event; $E(\cdot)$ denotes the expected value of a random variable; $Var(\cdot)$ denotes the variance; $Std(\cdot)$ denotes the standard deviation.

A. One-hop Sensor Localization Using RSS Measurements

In a two-dimensional plane, consider a single sensor (or source, target) located at the origin and N distance (or angle) measurements made to this sensor at N known locations, as illustrated in Figure 1. Here, the N known locations are abstracted as anchors and are labeled $1, \cdots, N$ with the i-th anchor's location denoted by $\mathbf{s}_i = [x_i, y_i]^T$. The true distance between the sensor and the i-th anchor is denoted by $d_i = \|\mathbf{s}_i\|$. The true angle subtended at the sensor by the i-th anchor and the positive x-axis is denoted θ_i.

For a specific localization problem, the precise locations of the N anchors, i.e. $[x_i, y_i]^T$, are given in advance; pair-wise distance measurements $\{\hat{d}_i, i = 1, \cdots, N\}$ between the sensor and the anchors are made and obey certain error models. Then, the aim of single-hop sensor localization is finding an estimate of the true sensor position using the observable set of distance measurements $\{\hat{d}_i, i = 1, \cdots, N\}$. In this paper, we consider the performance limit of sensor localization over a family of random anchor locations other than a specific localization problem with given anchor locations.

Let the sensor be a transmitter and the N anchors be receivers. Define $\{P_i, i = 1, \cdots, N\}$ to be the measured received signal powers at the N anchors transmitted by the sensor. We make the following assumptions:

Assumption 1: The wireless channel satisfies the log-normal (shadowing) model and the received powers $\{P_i, i = 1, \cdots, N\}$ at the N anchors are statistically independent.

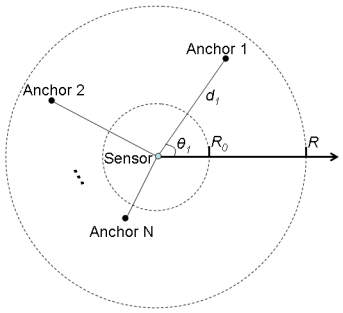


Fig. 1. Localizing a sensor using N anchors.

Remark 1: Assumption 1 is the basis for converting the RSS measurements (i.e. received powers) to distance estimates [5], and is commonly made in studies on RSS-based sensor localization (e.g. [1], [6]). It follows that $P_i(\text{dBm}) = 10 \log_{10} P_i$ are Gaussian

$$P_i(\text{dBm}) = \overline{P_0}(\text{dBm}) - 10\alpha \log_{10} \frac{d_i}{R_0} + Z, \qquad (1)$$

where $\overline{P_0}(\text{dBm})$ is the mean received power in dBm at a reference distance R_0, α is the path-loss exponent, and Z is a random variable representing the shadowing effect, normally distributed with mean zero and variance σ_{dB}^2 (in dBm). As pointed out in [7], due to the fact that the log-normal model does not hold for $d_i = 0$, the close-in distance R_0 is introduced as the known received power reference point, and is virtually the lower bound on practical distances used in the wireless communication system. Further, $\overline{P_0}(\text{dBm})$ is computed from the free space path loss formula (see, e.g. [7]).

B. A Random Sensor-Anchor Geometry Model

Assumption 2: The N anchors are randomly and uniformly distributed inside the annulus centered at the sensor and defined by radii R_0 and R ($R > R_0 > 0$).

Remark 2: In Assumption 2, R is the upper bound on practical distances which is normally restricted by the factors determining path loss attenuations; R_0, though representing the lower bound, is mainly devised to avoid the inconvenience in calculations, and theoretically speaking, any arbitrarily small positive number can be the lower bound. By Assumption 2, each possible sensor-anchor geometry is as probable as another, in the sense that the sensor-anchor geometry follows a "uniform" distribution. Furthermore, it is easy to show that $\{d_i, i = 1, \cdots, N\}$ and $\{\theta_i, i = 1, \cdots, N\}$ are mutually independent.

C. The Scalar Metric

The probability density function (pdf) of P_i can be formulated as follows

$$f_P(P_i) = \frac{10}{(\ln 10)\sqrt{2\pi}\sigma_{dB} P_i} \exp\left\{-\frac{b}{2}\left(\ln \frac{d_i}{\tilde{d}_i}\right)^2\right\}, \qquad (2)$$

where $b = \left(\frac{10\alpha}{\sigma_{dB} \ln 10}\right)^2$ and $\tilde{d}_i = d_0 \left(\frac{\overline{P_0}}{P_i}\right)^{\frac{1}{\alpha}}$.

For the purpose of computing the CRLB for sensor localization using the RSS measurements, we formulate the Fisher information matrix (FIM) F_{RSS} as

$$F_{RSS} = b \begin{pmatrix} \sum_{i=1}^{N} \frac{\cos^2\theta_i}{d_i^2} & \sum_{i=1}^{N} \frac{\cos\theta_i \sin\theta_i}{d_i^2} \\ \sum_{i=1}^{N} \frac{\cos\theta_i \sin\theta_i}{d_i^2} & \sum_{i=1}^{N} \frac{\sin^2\theta_i}{d_i^2} \end{pmatrix}. \quad (3)$$

A detailed derivation can be found in [1]. If F_{RSS} is non-singular, the CRLB, denoted C_{RSS}, is just the inverse of F_{RSS}. Then, we define $Tr(C_{RSS})$ to be a metric for the performance limit of localizing the sensor and have

$$Tr(C_{RSS}) = \frac{1}{b} \left(\frac{\sum_{i=1}^{N} \frac{1}{d_i^2}}{\sum_{1\le i<j\le N} \frac{\sin^2(\theta_i-\theta_j)}{d_i^2 d_j^2}} \right). \quad (4)$$

Since $\{d_i, i = 1, \cdots, N\}$ and $\{\theta_i, i = 1, \cdots, N\}$ are random variables, $Tr(C_{RSS})$ is obviously a random variable.

D. U-statistics

U-statistics are very natural in statistical work, particularly in the context of independent and identically distributed (i.i.d.) random variables, or more generally for exchangeable sequences, such as in simple random sampling from a finite population. The origins of the U-statistics theory are traceable to the seminal paper [4], which proved the Central Limit Theorems for U-statistics. Following the publication of this seminal paper, the interest in this class of statistics steadily increased, crystallizing into a well-defined and vigorously developing line of research in probability theory. Its formal definition is presented as follows:

Definition 1: Let $\{X_i, i = 1, \cdots, N\}$ be i.i.d. p-dimensional random vectors. Let $h(x_1, \cdots, x_r)$ be a Borel function on $\mathbb{R}^{r\times p}$ for a given positive integer r ($\le N$) and be symmetric in its arguments. A U-statistic U_N is

$$U_N = \frac{r!(N-r)!}{N!} \sum_{1\le i_1<\cdots<i_r\le N} h(X_{i_1}, \cdots, X_{i_r}) \quad (5)$$

and $h(x_1, \cdots, x_r)$ is called the kernel of U_N.

It is obvious that $Tr(C_{RSS})$ involves the ratio of two U-statistics according to (4), which motivates us to study $Tr(C_{RSS})$ through an asymptotic analysis based on the theory of U-statistics.

III. Main Results

Due to the complexity of $Tr(C_{RSS})$, it is very difficult to give its accurate distribution directly. As such, we endeavor to present an asymptotic analysis at first. Due to the space limit, proofs are omitted.

A. Theories

According to (4), a key property of $Tr(C_{RSS})$ is that it is the ratio of two sums of random variables, which can be processed by using the following lemma.

Lemma 1: Given $\{X_i^{(1)}, i = 1, \cdots, N\}$ and $\{X_i^{(2)}, i = 1, \cdots, N\}$ where

- $\{X_i^{(1)}, i = 1, \cdots, N\}$ are i.i.d. random variables with bounded values;
- $\{X_i^{(2)}, i = 1, \cdots, N\}$ are i.i.d. random variables with bounded values;
- $\{X_i^{(1)}, i = 1, \cdots, N\}$ and $\{X_i^{(2)}, i = 1, \cdots, N\}$ are mutually independent,

define vectors $X_i = [\ X_i^{(1)} \quad X_i^{(2)}\ ]^T$ $(i = 1, \cdots, N)$ and two sequences of random variables

$$T_N = \frac{1}{N}\sum_{i=1}^{N} X_i^{(1)}, \quad (6)$$

$$S_N = \frac{2}{N(N-1)} \sum_{1\le i<j\le N} \Big[X_i^{(1)} X_j^{(1)} \times \sin^2(X_i^{(2)} - X_j^{(2)}) \Big]. \quad (7)$$

Then, as $N \to \infty$,

$$\frac{T_N}{S_N} = \frac{1}{m_1 m_2} + \frac{2\sigma_1^2}{N m_1^3 m_2} + M_N + R_N \quad (8)$$

where $m_1 = E(X_1^{(1)})$, $\sigma_1 = Std(X_1^{(1)})$, $m_2 = E(\sin^2(X_1^{(2)} - X_2^{(2)}))$,

$$M_N = \frac{2}{N}\sum_{i=1}^{N} g_1(X_i) + \frac{2}{N(N-1)} \sum_{1\le i<j\le N} g_2(X_i, X_j), \quad (9)$$

$$g_1(X_i) = \frac{m_1 - X_i^{(1)}}{2m_1^2 m_2}, \quad (10)$$

$$g_2(X_i, X_j) = \frac{1}{m_1 m_2} - \frac{X_i^{(1)} + X_j^{(1)}}{m_1^2 m_2} + \frac{2X_i^{(1)} X_j^{(1)}}{m_1^3 m_2} - \frac{X_i^{(1)} X_j^{(1)} \sin^2(X_i^{(2)} - X_j^{(2)})}{m_1^3 m_2^2}, \quad (11)$$

and R_N is the remainder term. For any $\varepsilon > 0$, R_N satisfies

$$Pr\{|NR_N| \ge \varepsilon\} = O(N^{-1}), \quad (12)$$

$$Pr\{|N(\ln N)R_N| \ge \varepsilon\} = o(1), \quad (13)$$

In Lemma 1, by letting $X_i^{(1)} = \frac{1}{d_i^2}$ and $X_i^{(2)} = \theta_i$, we have $m_2 = 0.5$, and

$$m_1 = 2\left(\frac{\ln\frac{R}{R_0}}{R^2 - R_0^2}\right), \quad (14)$$

$$\sigma_1 = \sqrt{\frac{1}{R_0^2 R^2} - \left(\frac{2\ln\frac{R}{R_0}}{R^2 - R_0^2}\right)^2}, \quad (15)$$

and our main result is further summarized as follows.

Theorem 1: Let m_1 and σ_1 be defined by (14) and (15). Define a sequence of random variables

$$W_N = \left(\frac{\sqrt{N}(N-1)bm_1^2}{4\sigma_1}\right) Tr(C_{RSS}) - \frac{\sqrt{N}m_1}{\sigma_1} - \frac{2\sigma_1}{\sqrt{N}m_1}. \quad (16)$$

Then, as $N \to \infty$, W_N converges in distribution to a standard normal random variable.

Remark 3: In view of the linear relationship between W_N and $Tr(C_{RSS})$, it is clear that $Tr(C_{RSS})$ is asymptotically

normal. Therefore, for a sufficiently large N, the distribution of $Tr(C_{RSS})$ can be approximated by the normal distribution

$$\mathcal{N}\left(\frac{4}{(N-1)bm_1}\left(1+\frac{2\sigma_1^2}{Nm_1^2}\right),\left(\frac{4\sigma_1}{\sqrt{N}(N-1)bm_1^2}\right)^2\right). \quad (17)$$

Most importantly, the above normal random variable makes it possible for us to analytically study the performance limit, i.e. $Tr(C_{RSS})$. Firstly, we can obtain a comprehensive knowledge about how $Tr(C_{RSS})$ is statistically distributed and how $Tr(C_{RSS})$ is affected by N. Secondly, using the normal distribution function from (17), we can compute the probability that $Tr(C_{RSS})$ is below a given threshold for a known value of N; in turn, we can determine a threshold such that $Tr(C_{RSS})$ is below the threshold with a certain confidence level, say 0.99; in addition, we can find the minimum N such that $Tr(C_{RSS})$ is below a given threshold with a certain confidence level. Such analysis is undoubtedly helpful for the design and deployment of sensor networks. Thirdly, the moments of $Tr(C_{RSS})$ can be approximated by the corresponding moments of the normal variable defined by (17), namely,

$$E(Tr(C_{RSS})) \approx \frac{4}{(N-1)bm_1}+\frac{8\sigma_1^2}{N(N-1)bm_1^3}, \quad (18)$$

$$Std(Tr(C_{RSS})) \approx \frac{4\sigma_1}{\sqrt{N}(N-1)bm_1^2}, \quad (19)$$

which characterize the relationship among the mean and standard deviation of $Tr(C_{RSS})$, the number of anchors, noise statistics of the RSS measurements and the spatial distributions of the anchors.

A natural question arises as to how large N should be to obtain a good approximation; this gives rise to the convergence rate study. In the literature of U-statistics, the Berry-Esseen bound was developed for characterizing the convergence rates of U-statistics [8], [9]. Because W_N is affine to a U-statistic (i.e. M_N), we propose the following theorem describing the convergence rate of W_N in the way similar to the Berry-Esseen bound.

Theorem 2: Use the notations in Theorem 1 and define

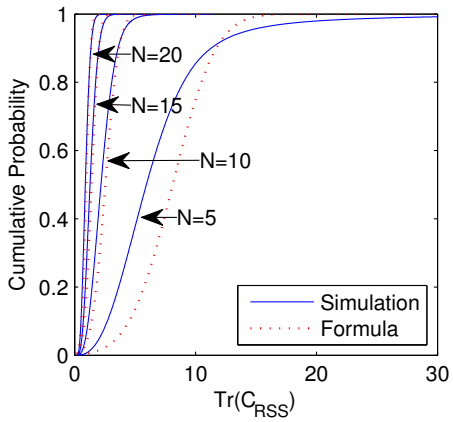
$$\nu_3 = E\left(\left(\frac{1}{d_1^2}-m_1\right)^3\right). \quad (20)$$

Then, as $N\to\infty$,

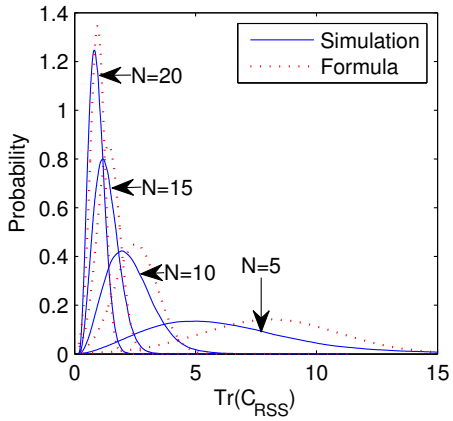
$$\sup_x |F_N(x)-\Phi(x)| \le \left|\left(\frac{\nu_3+\frac{2\sigma_1^4}{m_1}}{6\sigma_1^3}\right)\frac{(x^2-1)e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\right| N^{-\frac{1}{2}}+O(N^{-1}) \quad (21)$$

where $F_N(x)$ is the distribution function of W_N and $\Phi(x)$ is the standard normal distribution function.

Remark 4: Theorem 2 shows that as $N\to\infty$, the density of W_N converges to standard normality with the rate $O(N^{-\frac{1}{2}})$. Additionally, it can be verified that the coefficient associated with $N^{-\frac{1}{2}}$ is a function of the ratio $\frac{R}{R_0}$; that is to say, the convergence rate of the density of W_N is not determined by the individual values of R_0 and R, but by the ratio $\frac{R}{R_0}$.



(a)



(b)

Fig. 2. The distribution functions and pdfs of $Tr(C_{RSS})$ with $R_0 = 1\text{m}$, $R = 10\text{m}$, $\alpha = 2.3$ and $\sigma_{dB} = 3.92$.

IV. Simulations

In this subsection, we would like to carry out simulations to verify Theorem 1 and the approximations given in (18) and (19). The parameters α, σ_{dB} and R_0 describing the wireless channel are set as 2.3, 3.92 and 1 m, respectively, which are measured in a practical environment [1].

Firstly, we plot in Fig. 2(a) the actual distribution functions of $Tr(C_{RSS})$ (with the legend "Simulation") and the normal distribution function (17) (with the legend "Formula") for $N = 5, 10, 15, 20$. As can be seen, when $N = 5$, the discrepancy between them is quite obvious; when $N = 10$, the discrepancy becomes very small; when $N = 15$ or 20, the discrepancy is negligible. The discrepancy reduces with N increasing as illustrated in Fig. 2(a), and arises for two reasons: the intrinsic error in approximating a U-statistic by normality, and the existence of the remainder term R_N which obeys $Pr\left\{|R_N| \ge \frac{\varepsilon}{N}\right\} = O(N^{-1})$, see (12), and though nonzero is neglected in the calculation. Furthermore, we plot the corresponding pdfs in Fig. 2(b). It can be seen that the overall shapes of the actual pdfs (with the legend

"Simulation") are quite similar to those of normality, and the discrepancy in between reduces with N increasing. These observations are consistent with and in turn demonstrate Theorem 1.

Secondly, we plot the means and the standard deviations of $Tr(C_{RSS})$ from both simulations and the formulas (18) and (19) in Figs. 3(a), 3(c), 3(b) and 3(d). It is evident that the larger is N, the more precise are the formulas. When $N = 5$, the standard deviation attains comparatively large values, and the associated surface in Fig. 3(b) is non-smooth; the most probable reason is that the actual standard deviation is infinite for a N as small as 5.

For better comparison, we define the *relative error* to be the ratio of the difference between the quantity from the simulations and that from the corresponding formula to the former one, and plot them in Fig. 3(e) and 3(f). It can be seen that: (i) the mean is underestimated by (18) when R is small, say $R = 2$m, but is overestimated by (18) when R is large, say $R = 10$m, while the associated absolute value of the relative error decreases with N increasing in most cases; (ii) the standard deviation of $Tr(C_{RSS})$ is underestimated by (19) and the associated absolute value of the relative error decreases with increasing N and R; (iii) suppose the absolute value of the relative error below 10% is acceptable: when $R = 2$m, (18) is applicable if $N \geq 6$, but (19) is not applicable even if $N = 20$; when $R = 10$m, both (18) and (19) are applicable if $N \geq 11$.

In what follows, we present some useful remarks on the properties of sensor localization provided that (18) and (19) are applicable. It is notable that in (18) and (19), the mean and standard deviation of $Tr(C_{RSS})$ normalized by R^2 (or R_0^2) are dependent upon the ratio $\frac{R}{R_0}$; hence, we simplify the discussion involving R_0 and R by letting $R_0 = 1$ m and only concentrating on R.

Remark 5: Equation (18) quantitatively characterizes the average performance limit over all possible sensor-anchor geometries and is indicative for evaluating the average localization performance over a period of time and/or in a wide region. In addition, because the mean is in inverse proportion to N, a *critical* value N^* differing from the parameters R_0, R, σ_{dB} and α can be determined, such that having more anchors than N^* contributes little to the quality of sensor localization.

Remark 6: It can be easily deduced that both (18) and (19) monotonically decrease with R decreasing, as illustrated in Fig. 3(c) and 3(d); the reason is that long distance measurements from RSS suffer greater errors, and thus produce worse localization performance. Therefore, given a fixed N, distance measurements from a sensor are better made at locations as close to the sensor as possible. Moreover, it turns out that using more distance measurements spread over a wide range is not necessarily better than using fewer distance measurements but spread in a narrow range in terms of the average performance limit. For instance, $E(Tr(C_{RSS}))$ is approximately 0.52431 m^2 given $N = 15$ and $R = 6$ m, while a smaller mean which is approximately 0.43174 m^2 can be achieved given $N = 10$ and $R = 4$ m. Thus, tradeoff should be made between the number of anchors (i.e. N) and their spreading (i.e. R_0 and R) in sensor localization.

Remark 7: Though we discuss the impacts of N and R separately, the variables are correlated in some situations, and so the impacts are related. Normally, increasing all the transmission powers in a wireless sensor network enlarges the communication coverage of every node, and both N and R for localizing one sensor tend to rise, but $Tr(C_{RSS})$ and its mean will definitely decrease according to [1].

Remark 8: The dispersion of $Tr(C_{RSS})$ reflects its sensitivity to sensor-anchor geometries. Specifically, with a large dispersion, the chance of having two different sensor-anchor geometries to lead to a big difference in the resulting values of $Tr(C_{RSS})$ is large, implying a large sensitivity, and we should be careful about sensor-anchor geometries; by contrast, with a small dispersion, the chance is certainly small, so is the sensitivity, and there is less reason to worry about sensor-anchor geometries even if the anchors are randomly deployed. Given a random variable, the coefficient of variation, defined to be the ratio of its standard deviation to its mean, is a normalized measure of dispersion of its distribution. Therefore, the coefficient associated with $Tr(C_{RSS})$ has the order of $O(N^{-\frac{1}{2}})$ and the less is the coefficient, the smaller is the sensitivity. In particular, if the coefficient equals its minimum, i.e. 0, all the sensor-anchor geometries will result in one unique value of $Tr(C_{RSS})$, so that the minimum sensitivity is attained. Alternatively, we can observe the sensitivity from Fig. 2(b): the range of $Tr(C_{RSS})$ with a non-trivial probability becomes narrower and narrower with N increasing, implying that the sensitivity is reducing.

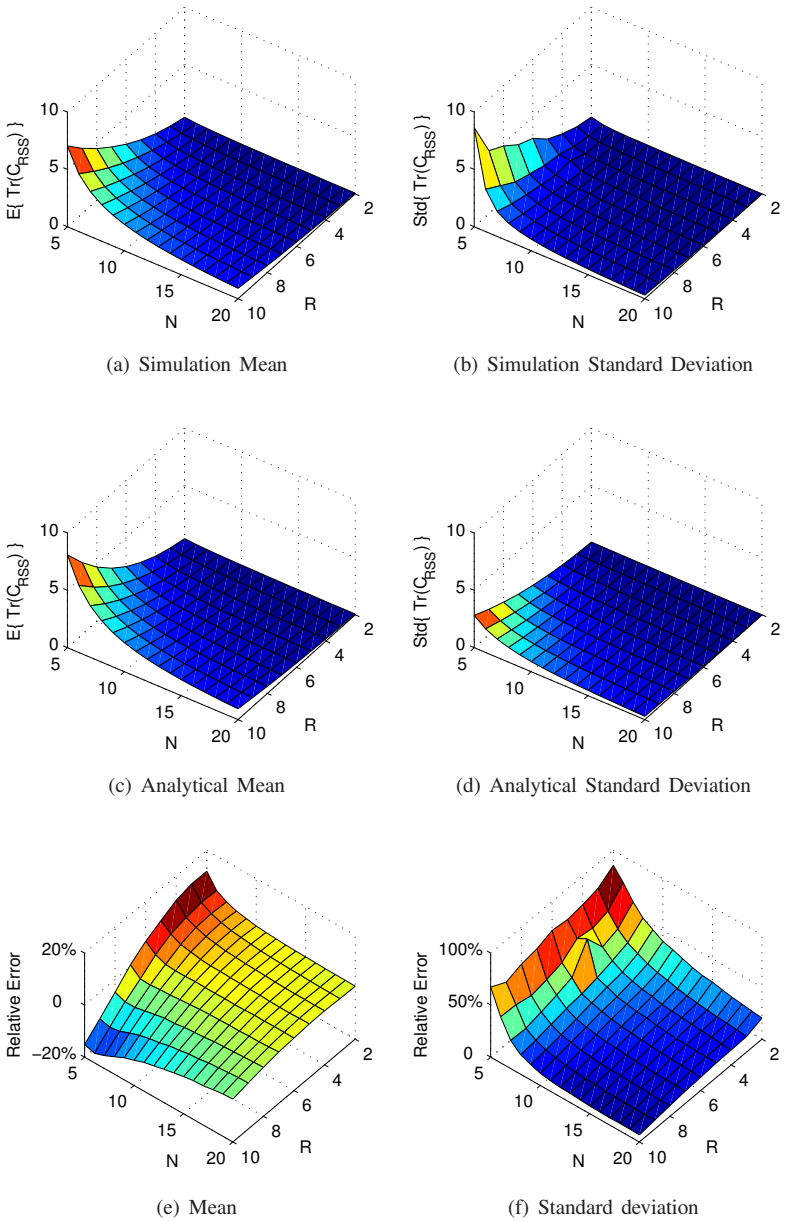
V. Conclusion and Future Work

In this paper, we investigated the performance limit of single-hop sensor localization with the RSS measurements by statistically sensor-anchor geometry modeling. That is, the positions of anchors are assumed to be random and the statistical attributes of the trace of the CRLB matrix embodies essential features of sensor localization. With strict mathematical proofs, we showed that the trace of the CRLB matrix is asymptotically normal. Based on this study, we analyzed the features of sensor localization and carried out extensive simulations.

In future work, we would like to take into account other distributions of anchor positions other than the uniform distributions, as well as considering other types of measuring techniques, including Time of Arrival (TOA), Time Difference of Arrival (TDOA), etc. In addition, it is more attractive, but of course extremely difficult, to conduct similar studies for multi-hop sensor localization.

Acknowledgment

B. Huang, C. Yu and B.D.O. Anderson are supported by the ARC (Australian Research Council) under DP-110100538 and NICTA. Tao Li is supported by the National Natural Science Foundation of China under grant 61004029. C. Yu is an ARC Queen Elizabeth II Fellow and



(a) Simulation Mean (b) Simulation Standard Deviation

(c) Analytical Mean (d) Analytical Standard Deviation

(e) Mean (f) Standard deviation

Fig. 3. The means and the standard deviations of $Tr(C_{RSS})$ from the simulations and the formulas, and the corresponding relative errors with $R_0 = 1\text{m}, \alpha = 2.3$ and $\sigma_{dB} = 3.92$.

is also supported by Overseas Expert Program of Shandong Province. B. Huang is also supported by the Key Laboratory of Computer Networks of Shandong Province. This material is based on research sponsored by the Air Force Research Laboratory, under agreement number FA2386-10-1-4102.

REFERENCES

[1] N. Patwari, A.O. III Hero, M. Perkins, N.S. Correal, and R.J. O'Dea. Relative location estimation in wireless sensor networks. *IEEE Transactions on Signal Processing*, 51(8):2137 – 2148, Aug. 2003.

[2] J.J. Leonard and H.F. Durrant-Whyte. Simultaneous map building and localization for an autonomous mobile robot. In *IEEE/RSJ International Workshop on Intelligent Robots and Systems.*, pages 1442–1447, Nov. 1991.

[3] P.N. Pathirana, N. Bulusu, A.V. Savkin, and S. Jha. Node localization using mobile robots in delay-tolerant sensor networks. *IEEE Transactions on Mobile Computing*, 4(3):285–296, May-Jun. 2005.

[4] W. Hoeffding. A class of statistics with asymptotically normal distribution. *The Annals of Mathematical Statistics*, 19(3):293–325, 1948.

[5] S.D. Chitte, S. Dasgupta, and Z. Ding. Distance estimation from received signal strength under log-normal shadowing: Bias and variance. *IEEE Signal Processing Letters*, 16(3):216–218, Mar. 2009.

[6] Y.-Y. Cheng and Y.-Y. Lin. A new received signal strength based location estimation scheme for wireless sensor network. *IEEE Transactions on Consumer Electronics*, 55(3):1295–1299, Aug. 2009.

[7] T. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall PTR, 2001.

[8] W.F. Grams and R.J. Serfling. Convergence rates for u-statistics and related statistics. *The Annals of Statistics*, 1(1):153–160, 1973.

[9] H. Callaert, P. Janssen, and N. Veraverbeke. An edgeworth expansion for u-statistics. *The Annals of Statistics*, 8(2):299–312, 1980.

Signal Processing 92 (2012) 427–438

Contents lists available at SciVerse ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Analyzing localization errors in one-dimensional sensor networks

Baoqi Huang [a,b], Changbin Yu [a,*], Brian D.O. Anderson [a,b]

[a] RSISE Building 115, The Australian National University, Canberra ACT 0200, Australia
[b] National ICT Australia Ltd., Australia

ARTICLE INFO

Article history:
Received 4 February 2011
Received in revised form
8 June 2011
Accepted 15 August 2011
Available online 23 August 2011

Keywords:
One-dimensional sensor networks
Localization
Estimation errors
Cramér–Rao lower bound
Error propagation

ABSTRACT

One-dimensional sensor networks can be found in many fields and demand node location information for various applications. Developing localization algorithms in one-dimensional sensor networks is trivial, due to the fact that existing localization algorithms developed for two- and three-dimensional sensor networks are applicable; nevertheless, analyzing the corresponding localization errors is non-trivial at all, because it is helpful to improving localization accuracy and designing sensor network applications. This paper deals with localization errors in distance-based multi-hop localization procedures of one-dimensional sensor networks through the Cramér–Rao lower bound (CRLB). We analyze the fundamental behaviors of localization errors and show that the localization error for a sensor is locally determined by network elements within a certain range of this sensor. Moreover, we break down the analysis of localization errors in a large-scale sensor network into the analysis in small-scale sensor networks, termed unit networks, in which tight upper and lower bounds on the CRLB can be established. Finally, we investigate two practical issues: the applicability of the analysis based on the CRLB and the optimal anchor placement.

1. Introduction

One-dimensional sensor networks can be found in many practical scenarios, such as high-voltage power lines [1–3], gas (water, oil) pipelines [4–7], mining tunnels [8], vehicular networks along a highway [9], bushfire detection [10] and so on. Regardless the dimensions of sensor networks, knowledge of node location is often used to report the geographic origin of events, to assist in target tracking, to achieve geographic aware routing, to manage sensor networks, to evaluate their coverage, and so on. Node localization is often posed as an optimization problem to estimate a most probable location among a set of possibilities [11–14], because distance and angle measurements suffer environmental and intrinsic hardware noises in reality and are then imprecise. Most of the existing localization algorithms, though developed for two- and three-dimensional sensor networks, can be directly applied in one-dimensional cases.

Localization algorithms usually assume the existence of anchors, i.e. nodes whose locations are known a priori or which are equipped with extra devices, such as GPS receivers, through which their locations can be obtained. The remaining nodes, termed sensors, besides being able to sense an event of interest typically can sense a distance or angle assisting in localization. If a sensor has direct measurements from a sufficient number of anchors to determine its location, its localization is termed one-hop localization; otherwise, it is termed multi-hop localization [15–17]. Localization using GPS is a specific case of one-hop localization; given a sensor network involving a small fraction of nodes as anchors, localization is normally multi-hop.

A challenging problem in relation to localization is analyzing localization errors and has received considerable

* Corresponding author. Tel.: +61 2 61258670; fax: +61 2 61258660.
E-mail addresses: baoqi.huang@anu.edu.au (B. Huang), brad.yu@anu.edu.au (C. Yu), brian.anderson@anu.edu.au (B.D.O. Anderson).

0165-1684/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2011.08.007

attention. For the one-hop case, localization errors are formulated by GDOP [18,19], the Cramér–Rao lower bound (CRLB), [15,16,20–22], Bayesian bounds [23] etc., and have been well understood; for the multi-hop case, on account of the complexity in the construction of localization errors, only simulation studies and very limited analytical studies have been introduced [24–26].

A closely relevant problem is known as "error propagation" [12,15]. Particularly, consider a two-dimensional multi-hop localization procedure using trilateration [14], in which a sensor is localizable only if it measures at least three distances from non-collinear anchors and/or already localized sensors (which perform as pseudo-anchors); if pseudo-anchors are involved, the errors in pseudo-anchor location estimates are propagated into location estimates of later localized sensors. In effect, the phenomenon of propagating (i.e. growing) localization errors is a common characteristic in the case of multi-hop localization. In view of the increasing deployment of sensor networks for various applications, it is evidently desirable to understand how localization errors grow among sensors given a configuration of sensor network parameters, such as sensor density, anchor density, anchor placement strategy, level of distance measurement errors, etc. Based on the study of localization errors, developers can choose proper parameters to control localization errors; in addition, location estimates provide context to sensed data (e.g. temperature, humidity, etc), and the study of localization errors underpins assessment of the reliability of location estimates of different sensors, which is helpful to further data processing in various sensor network applications.

In this paper, we attempt to analyze localization errors in distance-based multi-hop localization procedures of one-dimensional sensor networks through the CRLB. The usage of the CRLB is because it is algorithm-independent and can be attained by the localization algorithm using the maximum-likelihood estimator (MLE) in the one-dimensional case (as is proved in this paper). By treating distance measurements between nodes as connections constituting networks, we consider a sensor network as a general network. Because of the complexity associated with a general network, it proves useful to consider by way of an intermediate step *unit networks*, which are defined as follows:

Definition 1. A one-dimensional connected network is a unit network if and only if all anchors are placed at the left side of all sensors and/or the right side of all sensors.

As shown in Fig. 1, there are no anchors between any pair of sensors in unit networks. In the light of this feature, a divide and conquer method is developed in the sense that by building on results for unit networks, conclusions regarding general networks can be drawn, as will be demonstrated.

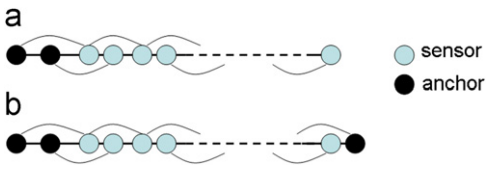


Fig. 1. Examples of unit networks. In (a), anchors are placed in the leftmost side; in (b), anchors are placed in both the leftmost and the rightmost sides.

The major contributions of this paper include: proofs for the fundamental behavior of localization errors in one-dimensional sensor networks; a divide and conquer method for analyzing localization errors in large-scale one-dimensional sensor networks; the behavior of localization errors in small-scale unit networks; the consideration of two practical issues regarding the analysis, viz. applicability of the analysis based on the CRLB and the optimal anchor placement.

The remainder of this paper is organized as follows. The next section reviews the literature. Section 3 provides the formulation of the CRLB in one-dimensional networks and presents certain fundamental behaviors of localization errors. Section 4 proposes a divide and conquer method to analyze localization errors in large-scale networks based on unit networks. Section 5 further discusses two practical issues. Finally, Section 6 concludes the paper and sheds light on future work.

2. Related work

In this section, we first review the algorithm-independent analysis of localization errors and then introduce some analysis focusing on error propagation.

In [15], Savvides et al. investigated the different aspects of the error of location estimates in distance-based multi-hop localization procedures. They formulated the CRLBs for sensors in a network under the Gaussian model which assumes that the errors in distance measurements are i.i.d. (independent and identically distributed) Gaussian with mean zero, and computed the root mean squared error (RMSE) as a metric for the localization accuracy of this network using the derived CRLBs. Then, simulation was carried out to study the impacts of various network parameters, including sensor and anchor densities, level of distance measurement errors and network size, on the RMSE. Additionally, a similar study was reported for networks with distance and/or angle measurements in [16].

In [20], Chang et al. also dealt with distance-based multi-hop localization procedures under the Gaussian model, and provided a geometric interpretation of the CRLB in the sense that the CRLB is essentially invariant under zooming, translation, and rotation. Moreover, they derived lower and upper bounds on the CRLB in two-dimensional networks which are similar to the bounds we will establish in Section 4. Apart from this, they discussed at length the error behaviors of a special kind of localization procedures in which no anchors are involved and only relative location can be obtained.

In [23], Wang et al. focused on the error behaviors in one-hop localization procedure. Given anchor locations and errors in distance measurements from a sensor to the anchors, their method computes the minimum-entropy location distribution of this sensor. They defined the Bayesian bound as the lower bound on the covariance of such distribution, which was compared with the CRLB

through simulations. Particularly, under the Gaussian model, the Bayesian bound equals the CRLB.

Formulations of the CRLB for localization problems were also presented in [21,22]: Patwari et al. derived the CRLB under Gaussian and log-normal models respectively in [21]; Larsson included clock biases existing in TOA (time of arrival) measurement systems as unknown parameters in the formulation of the CRLB in [22].

Besides these studies on the general behavior of localization errors, very limited results were also derived for error propagation. Niculescu et al. in [24] provided a closed-form formula describing error propagation in terms of the fraction of anchors, node density and the level of measurement errors. Their result is not valid for purely distance-based localization algorithms but for a particular localization algorithm, termed DV-Position, which requires both distance and AOA (angle of arrival) measurements.

In [25], Shi et al. studied error propagation by performing a theoretical analysis for one-dimensional networks. Two major conclusions were reported: the localization accuracy improves by adding more sensors into the network; for multi-hop localization procedures, the localization accuracy of a sensor is mainly dependent upon the number of hops it is away from anchors, so that the further away from anchors is a sensor, the worse is its localization accuracy. For two- and three-dimensional cases, these conclusions were extended by simulations and experiments. Nevertheless, their theoretical treatments are not adequate because they are restricted to highly simple one-dimensional networks.

In addition, [26] formally analyzed the algorithmic complexity of localization using interferometric ranging and proposed an iterative algorithm that gradually determines sensor locations hop by hop in a multi-hop localization procedure. Given sensors randomly distributed within a square and four anchors deployed to the center of the square, simulations based on the iterative algorithm indicated a linear increase of the localization errors (measured by the average distance between true locations and the corresponding estimated locations) with the hop count[1] from a sensor to anchors, but theoretical explanation was not provided.

In summary, theoretically analyzing localization errors is still a challenging problem in the literature, and through this effort, we hope to present a theoretical understanding on how localization errors behave during the multi-hop localization procedures of one-dimensional networks.

3. Fundamental behaviors of localization errors

In this section, after formulating the CRLB in one-dimensional networks, we shall discuss in detail several factors affecting the CRLB as well as localization errors. We use the following mathematical notations throughout this paper: $E(\cdot)$ denotes the expected value; $Var(\cdot)$ denotes the variance; $(\cdot)^{-1}$ denotes the matrix inverse; $\mathbf{x}$ denotes a vector; X_n denotes a square matrix of order n; $(X)_{ij}$ or $(X)_{i,j}$ denotes the entry in the i-th row and j-th column of matrix X; $(X)_i$ denotes the i-th row of matrix X.

3.1. The problem model of one-dimensional networks

Define a one-dimensional network $\mathcal{N}$ to be a triple (A,S,M), where A denotes the anchor set, S the sensor set and M the distance measurement set. Assume that

- all nodes, including both anchors and sensors, are deployed along a straight line;
- two nodes can measure their distance if and only if it is less than 1;
- anchor locations are precisely known and sensor locations are unknown;
- errors arise in distance measurements which are independent Gaussian with mean zero and certain variances differing from the lengths of actual distances.

Construct an undirected graph $G=(V,E)$ for $\mathcal{N}$, where $V=A\cup S$ and $E=M$. If G is connected, $\mathcal{N}$ is connected; otherwise, $\mathcal{N}$ is in fact a group of networks with connected graphs and can be studied separately. Hence, we assume that in this paper G is connected, and say $\mathcal{N}$ is connected for simplicity.

3.2. CRLB

Given a one-dimensional network $\mathcal{N}=(A,S,M)$ conforming to the problem model, we need to formulate the Fisher information matrix (FIM) for $\mathcal{N}$ to compute its CRLB. Define the following notations:

- $n=|S|$ and sensors are labeled as $1,\ldots,n$;
- $m=|A|$ and anchors are labeled as $n+1,n+2,\ldots,n+m$;
- the true location of i $(1\le i\le n+m)$ is x_i;
- the true and noisy distance measurements between i and j are d_{ij} and $\overline{d_{ij}}$ $(i<j)$;
- p_{ij} is the probability density function of $\overline{d_{ij}}$;
- the error in $\overline{d_{ij}}$ is $e_{ij}=\overline{d_{ij}}-d_{ij}$ and its standard deviation is σ_{ij} (>0);
- J_n is an $n\times n$ square matrix and denotes the FIM for $\mathcal{N}$;
- $\emptyset$ denotes an empty set.

In terms of estimation terminologies, suppose $X=\{x_1,\ldots,x_n\}$ is the set of parameters to be estimated and M is the set of observations. Because of the independent Gaussian errors, the *logarithmic* likelihood function is

$$\ln f(M;X)=\sum_{\overline{d_{ij}}\in M}\ln p_{ij}, \tag{1}$$

$$p_{ij}=\frac{1}{\sqrt{2\pi}\sigma_{ij}}\exp\left\{-\frac{(\overline{d_{ij}}-|x_i-x_j|)^2}{2\sigma_{ij}^2}\right\}, \tag{2}$$

[1] Throughout this paper, "hop count" between two nodes refers to the number of hops along their shortest path.

where $i,j=1,\ldots,n+m$. We can obtain

$$(J_n)_{ij}=E\left(\frac{\partial}{\partial x_i}\ln f(M;X)\frac{\partial}{\partial x_j}\ln f(M;X)\right) \tag{3}$$

where $i,j=1,\ldots,n$ and then

$$(J_n)_{ij}=\begin{cases}\sum\limits_{\overline{d_{ik}}\in M}\frac{1}{\sigma_{ik}^2}+\sum\limits_{\overline{d_{ki}}\in M}\frac{1}{\sigma_{ki}^2}, & i=j,\\ -\frac{1}{\sigma_{ij}^2}, & \overline{d_{ij}}\in M,\\ -\frac{1}{\sigma_{ji}^2}, & \overline{d_{ji}}\in M,\\ 0 & \text{otherwise.}\end{cases}$$

Obviously, J_n is symmetric. Furthermore, if J_n is non-singular, we define

$$C_n=J_n^{-1}, \tag{4}$$

the diagonal entries of which are the CRLBs on the variances of localization errors. Theorem 1 provides a sufficient and necessary condition for the existence of C_n.

Theorem 1. *Given a connected network $\mathcal{N}=(A,S,M)$, its FIM J_n is positive-definite if and only if $A\neq\emptyset$.*

Proof. Define $\mathbf{y}=[y_1,\ldots,y_n]^T$ to be a column vector with n real entries. Then

$$\mathbf{y}^TJ_n\mathbf{y}=\sum_{\overline{d_{ij}}\in M}\frac{1}{\sigma_{ij}^2}(y_i-y_j)^2+\sum_{i=1}^{n}\left(\sum_{\overline{d_{ik}}\in M\wedge k\in A}\frac{1}{\sigma_{ik}^2}y_i^2\right). \tag{5}$$

Obviously, J_n is positive-semidefinite, and thus we only need to prove that J_n is positive-definite, which is equivalent to the statement that J_n is nonsingular, or that $y_1=\cdots=y_n=0$ is the unique solution to

$$\sum_{\overline{d_{ij}}\in M}\frac{1}{\sigma_{ij}^2}(y_i-y_j)^2+\sum_{i=1}^{n}\left(\sum_{\overline{d_{ik}}\in M\wedge k\in A}\frac{1}{\sigma_{ik}^2}y_i^2\right)=0. \tag{6}$$

Due to the non-negativity of the two terms, we obtain

$$\begin{cases}\sum\limits_{\overline{d_{ij}}\in M}\frac{1}{\sigma_{ij}^2}(y_i-y_j)^2=0,\\ \sum\limits_{i=1}^{n}\left(\sum\limits_{\overline{d_{ik}}\in M\wedge k\in A}\frac{1}{\sigma_{ik}^2}y_i^2\right)=0,\end{cases} \tag{7a, b}$$

Each sensor is associated with an entry of $\mathbf{y}$. From (7a), we know if two sensors are directly connected, the associated entries of $\mathbf{y}$ are equal; given two sensors, if a path consisting of only sensors exists between them, the associated entries of $\mathbf{y}$ are also equal; from (7b), we know if a sensor is directly connected to an anchor, its associated entry of $\mathbf{y}$ must be zero. Now if $A=\emptyset$, because $\mathcal{N}$ is connected, paths between any two sensors only consist of sensors and all entries of $\mathbf{y}$ are equal but their value is arbitrary, and hence J_n is only positive-semidefinite; on the other hand, if $A\neq\emptyset$, because $\mathcal{N}$ is connected, each sensor is either directly connected to an anchor or indirectly connected to an anchor through other sensors and thus the associated entry of $\mathbf{y}$ must be zero, and hence J_n is positive-definite. We conclude that when $\mathcal{N}$ is connected, J_n is positive-definite if and only if $A\neq\emptyset$. □

3.3. Factors affecting the CRLB

In the problem model, it is evident that distance measurement errors are the only source of localization errors. Apart from that, network topology related factors such as node degrees, connectivity, hop counts to anchors, etc., also affect the magnitude of localization errors. In this subsection, we shall discuss the effects of these factors on the CRLB by examining some primitive network modifications:

- adding an additional distance measurement between two existing sensors;
- adding an additional sensor and an additional distance measurement between itself and another sensor;
- adding an additional anchor and an additional distance measurement between itself and a sensor;
- replacing a sensor by an additional anchor.

From the network deployment point of view, the first two modifications correspond to increasing sensor density by adding more sensors or enlarging the range of distance measurements, and the last two modifications correspond to increasing anchor density. In what follows, Theorems 2–5 will describe their effects.

Theorem 2. *Suppose $\mathcal{N}=(A,S,M)$ $(A\neq\emptyset)$ is a connected network and its FIM is J_n. If there is no distance measurement between sensor i and j $(i<j)$ in $\mathcal{N}$, let $\tilde{M}=M\cup\{\overline{d_{ij}}\}$. Construct a new network $\tilde{\mathcal{N}}=(A,S,\tilde{M})$. Let $\tilde{J}_n$ be the FIM of $\tilde{\mathcal{N}}$ and C_n and $\tilde{C}_n$ be the inverses of J_n and $\tilde{J}_n$. Then, (1) for any sensor s $(1\leq s\leq n)$, $(\tilde{C}_n)_{ss}\leq(C_n)_{ss}$; (2) $\mathrm{Tr}(\tilde{C}_n)<\mathrm{Tr}(C_n)$.*

Proof. Suppose $\mathbf{y}$ is a column vector of order n with all entries 0 except for the i-th one being 1 and the j-th one being -1. Then, we have

$$\tilde{J}_n=J_n+\frac{1}{\sigma_{ij}^2}\mathbf{y}\mathbf{y}^T. \tag{8}$$

Using the method of inverting perturbed matrices, see e.g. [27], we obtain

$$\tilde{C}_n=C_n-\frac{((C_n)_i-(C_n)_j)^T((C_n)_i-(C_n)_j)}{\sigma_{ij}^2-2(C_n)_{ij}+(C_n)_{ii}+(C_n)_{jj}}, \tag{9}$$

$$(\tilde{C}_n)_{ss}=(C_n)_{ss}-\frac{((C_n)_{is}-(C_n)_{js})^2}{\sigma_{ij}^2-2(C_n)_{ij}+(C_n)_{ii}+(C_n)_{jj}}. \tag{10}$$

Since C_n is positive-definite, $|(C_n)_{ij}|<\sqrt{(C_n)_{ii}(C_n)_{jj}}$. This together with $(C_n)_{ii}+(C_n)_{jj}\geq 2\sqrt{(C_n)_{ii}(C_n)_{jj}}$ implies that $-2(C_n)_{ij}+(C_n)_{ii}+(C_n)_{jj}>0$ and thus $(\tilde{C}_n)_{ss}\leq(C_n)_{ss}$.

Moreover, define $\lambda_k(J_n)$ $(k=1,\ldots,n)$ to be the eigenvalues of J_n which are listed in decreasing order. Because $\lambda_1(\mathbf{y}\mathbf{y}^T)=2$ and $\lambda_2(\mathbf{y}\mathbf{y}^T)=\cdots=\lambda_n(\mathbf{y}\mathbf{y}^T)=0$, $\mathbf{y}\mathbf{y}^T$ is positive-semidefinite.

Because J_n is positive-definite, it is known [28] that

$$0 < \lambda_k(J_n) \le \lambda_k(\tilde{J}_n). \tag{11}$$

Due to $\mathrm{Tr}(\mathbf{y}\mathbf{y}^T) > 0$, $\mathrm{Tr}(\tilde{J}_n) > \mathrm{Tr}(J_n)$ and hence $\sum_{k=1}^{n} \lambda_k(\tilde{J}_n) > \sum_{k=1}^{n} \lambda_k(J_n)$; together with (11), it can be concluded that there exists $l \in 1,\ldots,n$ such that $\lambda_l(\tilde{J}_n) > \lambda_l(J_n)$; accordingly, there exists $\bar{l} \in 1,\ldots,n$ such that $\lambda_{\bar{l}}(\tilde{J}_n^{-1}) < \lambda_{\bar{l}}(J_n^{-1})$, which implies that $\mathrm{Tr}(\tilde{C}_n) < \mathrm{Tr}(C_n)$. □

Remark 1. Theorem 2 indicates that adding an additional distance measurement between two existing sensors either reduces the CRLB for any sensor or keeps it unchanged, while the total CRLB in the network is definitely reduced. Specifically, in light of (10), the CRLB for sensor s is unchanged if and only if $(C_n)_{is} = (C_n)_{js}$, namely that the correlation between location estimates of sensors i, s equals that associated with sensors j,s. As shown in Fig. 2(b), it is obvious that the CRLBs for sensors 2,4,5 do not change by adding the distance measurement between sensors 1 and 3. From the network topology point of view, a larger node degree, i.e. more distance measurements or a higher sensor density, corresponds to smaller CRLBs.

Remark 2. As a rule of thumb, for sensors i and j, the smaller is the difference between the hop counts from them to sensor s, usually the less is the difference of correlations, i.e. $|(C_n)_{is} - (C_n)_{js}|$, and thus the less is $(C_n)_{ss} - (\tilde{C}_n)_{ss}$ based on (10). Therefore, a distance measurement between two sensors with a large hop count results in a greater reduction in localization errors than that with a small hop count.

Next, we consider the introduction of an additional sensor.

Theorem 3. *Suppose $\mathcal{N} = (A,S,M)$ $(A \neq \emptyset)$ is a connected network with the FIM J_n and an additional sensor is labeled as $n+m+1$. Let $\tilde{S} = S \cup \{n+m+1\}$ and $\tilde{M} = M \cup \{\overline{d_{i,n+m+1}}\}$ $(1 \le i \le n)$. Construct a new network $\tilde{\mathcal{N}} = (A,\tilde{S},\tilde{M})$. Let $\tilde{J}_{n+1}$ be the FIM of $\tilde{\mathcal{N}}$ and C_n and $\tilde{C}_{n+1}$ be the inverses of J_n and $\tilde{J}_{n+1}$. Then, (1) the leading $n \times n$ submatrix of $\tilde{C}_{n+1}$ is identical to C_n; (2) $(\tilde{C}_{n+1})_{n+1,n+1} = (C_n)_{ii} + \sigma^2_{i,n+m+1}$.*

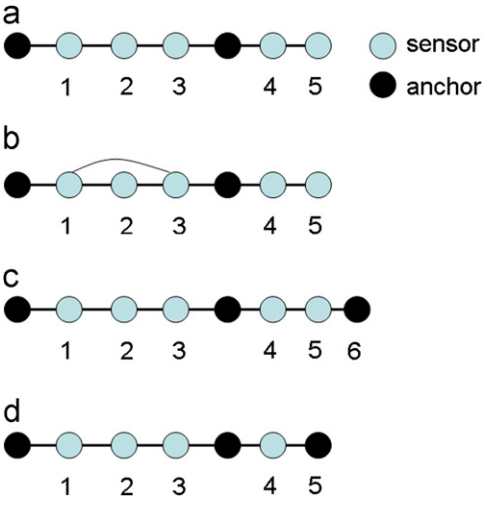


Fig. 2. A network and its variations based on different primitive modifications. Given the original network in (a), an additional distance measurement is added between sensors 1 and 3 in (b), an additional anchor 6 and an associated distance measurement are added in (c), and the sensor 5 is replaced by an anchor in (d).

Proof. Suppose $\mathbf{y}$ is a column vector of order n with all entries 0 except for the i-th one being 1, and according to the calculations of the FIM, we can obtain

$$\tilde{J}_{n+1} = \begin{pmatrix} J_n + \frac{1}{\sigma^2_{i,n+m+1}}\mathbf{y}\mathbf{y}^T & -\frac{1}{\sigma^2_{i,n+m+1}}\mathbf{y} \\ -\frac{1}{\sigma^2_{i,n+m+1}}\mathbf{y}^T & \frac{1}{\sigma^2_{i,n+m+1}}\mathbf{y}^T\mathbf{y} \end{pmatrix}. \tag{12}$$

By using block matrix inversion, we obtain

$$\tilde{J}^{-1}_{n+1} = \begin{pmatrix} J_n^{-1} & \cdots \\ \cdots & \mathbf{y}^T J_n^{-1}\mathbf{y} + \sigma^2_{i,n+m+1} \end{pmatrix}, \tag{13}$$

$$\tilde{C}_{n+1} = \begin{pmatrix} C_n & \cdots \\ \cdots & (C_n)_{ii} + \sigma^2_{i,n+m+1} \end{pmatrix}, \tag{14}$$

which proves the theorem. □

Remark 3. Theorem 3 shows that an additional sensor and an associated distance measurement do not change the CRLBs in the original network, and the CRLB for the additional sensor equals the sum of the variance in the distance measurement and the CRLB for the sensor that the distance measurement is from.

Then, we consider an additional distance measurement from an additional anchor.

Theorem 4. *Suppose $\mathcal{N} = (A,S,M)$ $(A \neq \emptyset)$ is a connected network with the FIM J_n and an additional anchor is labeled as $n+m+1$. Let $\tilde{A} = A \cup \{n+m+1\}$ and $\tilde{M} = M \cup \{\overline{d_{i,n+m+1}}\}$ $(1 \le i \le n)$. Construct a new network $\tilde{\mathcal{N}} = (\tilde{A},S,\tilde{M})$. Let $\tilde{J}_n$ be the FIM of $\tilde{\mathcal{N}}$ and C_n and $\tilde{C}_n$ be the inverses of J_n and $\tilde{J}_n$. Then, (1) for every sensor s $(1 \le s \le n)$, $(C_n)_{ss} \ge (\tilde{C}_n)_{ss}$; (2) $\mathrm{Tr}(\tilde{C}_n) < \mathrm{Tr}(C_n)$.*

Proof. Let $\mathbf{y}$ be the same vector as in the proof of Theorem 3. Based on the formulation of the FIM, we have

$$\tilde{J}_n = J_n + \frac{1}{\sigma^2_{i,n+m+1}}\mathbf{y}\mathbf{y}^T. \tag{15}$$

Using the *Sherman–Morrison–Woodbury formula*, see e.g. [28], we obtain

$$\tilde{C}_n = C_n - \frac{(C_n\mathbf{y})(C_n\mathbf{y})^T}{\sigma^2_{i,n+m+1} + \mathbf{y}^T C_n \mathbf{y}}, \tag{16}$$

$$(\tilde{C}_n)_{ss} = (C_n)_{ss} - \frac{(C_n)^2_{si}}{\sigma^2_{i,n+m+1} + (C_n)_{ii}}. \tag{17}$$

Due to $(C_n)_{ii} > 0$, we derive $(\tilde{C}_n)_{ss} \le (C_n)_{ss}$. By (15) and following the same argument as in the proof of Theorem 2, we can show that $\mathrm{Tr}(\tilde{C}_n) < \mathrm{Tr}(C_n)$. □

Remark 4. Theorem 4 indicates that introducing an additional distance measurement from an additional

anchor either reduces the CRLB for any sensor or keeps it unchanged, but reduces the total CRLB in the network. In view of (17), such modification does not change the CRLB for sensor s if and only if $(C_n)_{si}=0$, namely that there is no correlation between location estimates of sensors s,i. Fig. 2(c) illustrates an example, in which the CRLBs for sensors 1–3 are unchanged even if an additional anchor 6 and an associated distance measurement are introduced. Evidently, these conclusions are still correct, though the additional distance measurement is made from an existing anchor rather than from an additional anchor.

Remark 5. Provided that $s=i$ in (17), we can obtain $(\tilde{C}_n)_{ii}=(\sigma^2_{i,n+m+1}(C_n)_{ii}/(\sigma^2_{i,n+m+1}+(C_n)_{ii}))<\sigma^2_{i,n+m+1}$, which reveals that no matter how large $(C_n)_{ii}$ is, an additional distance measurement between sensor i and an anchor always makes $(\tilde{C}_n)_{ii}$ be less than $\sigma^2_{i,n+m+1}$. Interpreted from another point of view, the positions of anchors have great influence on the CRLBs in the sense that a small hop count of a sensor away from anchors probably leads to a small CRLB and thus a small error.

Remark 6. Based on (17), we define

$$r_{si}=\frac{(C_n)_{ss}-(\tilde{C}_n)_{ss}}{(C_n)_{ss}}=\frac{\rho^2_{si}}{1+\dfrac{\sigma^2_{i,n+m+1}}{(C_n)_{ii}}}, \tag{18}$$

where $i,s=1,\ldots n$ and ρ_{si} ($|\rho_{si}|\le 1$) denotes the correlation coefficient between i and s. Hence, r_{si} measures the relative reduction in the CRLB for sensor s. The coefficients ρ_{si} depend on network topologies and may require a lengthy calculation for their determinations. Simulations however reveal ρ_{si} reduces with the hop count between i and s increasing at a nearly linear rate. Thus, according to (18), the effect of an additional anchor dies off at a quadratic speed in a direction moving away from this anchor.

Remark 7. Given a network and an additional anchor, the optimal placement of this anchor will maximize $\sum_{s=1}^{n}((C_n)_{ss}-(\tilde{C}_n)_{ss})$, which, see (17), is equivalent to determining the value of i. The coefficient $|\rho_{si}|$ reaches the maximum 1 when $s=i$. Assume that ρ_{si} increases monotonically with s increasing from 0 to i and decreases monotonically from i to n. By observing (18), given a relatively small $(C_n)_{ii}$, $1/(1+\sigma^2_{i,n+m+1}/(C_n)_{ii})$ is relatively small, and with s around i, $(C_n)_{ss}$ is also small but ρ^2_{si} is comparatively large, which probably leads to a small sum; on the other hand, for a relatively large $(C_n)_{ii}$, $1/(1+\sigma^2_{i,n+m+1}/(C_n)_{ii})$ is relatively large, and with s around i, both $(C_n)_{ss}$ and ρ^2_{si} are comparatively large, which probably leads to a large sum. Thus, an additional anchor should be put near the sensor with the largest CRLB. Because sensors near anchors always have small CRLBs, additional anchors should not be placed near these sensors and hence anchors should not be clustered. Essentially, adding an additional anchor will probably result in more than one additional distance measurement. But if we take into account these distance measurements one by one, based on this remark, effects of other distance measurements may be trivial compared with that of the first one. Actually, it does not matter which one should be the first.

Lastly, we consider replacing an existing sensor by an additional anchor.

Theorem 5. *Suppose $\mathcal{N}=(A,S,M)$ $(A\neq\emptyset)$ is a connected network with the FIM J_n and $n\in S$. Let $\tilde{A}=A\cup\{n\}$ and $\tilde{S}=S-\{n\}$. Construct a new network $\tilde{\mathcal{N}}=(\tilde{A},\tilde{S},M)$. Let $\tilde{J}_{n-1}$ be the FIM of $\tilde{\mathcal{N}}$ and C_n and $\tilde{C}_{n-1}$ be the inverses of J_n and $\tilde{J}_{n-1}$. Then, (1) for every sensor s $(1\le s\le n-1)$, $(\tilde{C}_{n-1})_{ss}\le(C_n)_{ss}$; (2) $\mathrm{Tr}(\tilde{C}_{n-1})<\sum_{k=1}^{n-1}(C_n)_{kk}$.*

Proof. Let $\mathbf{y}$ be a column vector consisting of the leading $n-1$ entries in $(J_n)_n$. Based on the formulation of the FIM, we have

$$J_n=\begin{pmatrix}\tilde{J}_{n-1} & \mathbf{y}\\ \mathbf{y}^T & (J_n)_{nn}\end{pmatrix}, \tag{19}$$

$$C_n=\begin{pmatrix}(\tilde{J}_{n-1}-(J_n)_{nn}\mathbf{y}\mathbf{y}^T)^{-1} & \cdots\\ \cdots & \cdots\end{pmatrix}. \tag{20}$$

Suppose the leading $(n-1)\times(n-1)$ submatrix of C_n is V^{-1}_{n-1}, which is positive-definite since C_n is positive-definite. Then, we can obtain

$$\tilde{J}_{n-1}=V_{n-1}+(J_n)_{nn}\mathbf{y}\mathbf{y}^T, \tag{21}$$

$$\tilde{C}_{n-1}=V^{-1}_{n-1}-\frac{(J_n)_{nn}(V^{-1}_{n-1}\mathbf{y})(V^{-1}_{n-1}\mathbf{y})^T}{1+(J_n)_{nn}\mathbf{y}^TV^{-1}_{n-1}\mathbf{y}}, \tag{22}$$

$$(\tilde{C}_{n-1})_{ss}=(V^{-1}_{n-1})_{ss}-\frac{((V^{-1}_{n-1})_s\mathbf{y})^2}{\dfrac{1}{(J_n)_{nn}}+\mathbf{y}^TV^{-1}_{n-1}\mathbf{y}}. \tag{23}$$

Rewrite $\mathbf{y}^TV^{-1}_{n-1}\mathbf{y}$ to

$$\mathbf{y}^TV^{-1}_{n-1}\mathbf{y}=(\mathbf{y}^T\ \ 0)C_n\begin{pmatrix}\mathbf{y}\\0\end{pmatrix}. \tag{24}$$

Obviously, $\mathbf{y}^TV^{-1}_{n-1}\mathbf{y}\ge 0$. This with $(J_n)_{nn}>0$ implies $(\tilde{C}_{n-1})_{ss}\le(C_n)_{ss}$. By (21) and following the same argument in Theorem 2, we can prove the second part of this theorem. □

Remark 8. This theorem reveals that replacing a sensor in a network by an anchor and keeping all distance measurements unchanged either reduces the CRLB for any sensor or keeps it unchanged, but the total CRLB in the network (excluding the sensor being replaced) is definitely reduced. According to (23), the sufficient and necessary condition for the unchanged CRLB is that $(V^{-1}_{n-1})_s\mathbf{y}$ equals 0. Fig. 2(d) shows an example in which the CRLBs for sensors 1,2,3 are unchanged after sensor 5 is replaced by an anchor.

4. A divide and conquer analysis in large-scale sensor networks

A common view is that the localization error of a sensor is mainly determined by nearby network elements, while those distant sensors and anchors have little effects. In this section, we shall theoretically analyze this point

and then propose a divide and conquer method to analyze localization errors in large-scale networks.

4.1. Simulation setup

In this and subsequent sections, we shall carry out extensive simulations in Matlab to verify our conclusions. One common step involved in all simulations is generating one-dimensional networks: a network is produced by randomly and uniformly distributing n sensors in the range of $[0, L]$. A network is called a valid network only if it is connected and the locations 0 and L are within the distance measurement scopes of the leftmost and the rightmost sensors respectively. If a network is invalid, we will repeat until it is valid. For simplicity, we assume that the standard deviations of distance measurement errors are identically 1, i.e. $\sigma_{ij}=1$. The values of n and L will be assigned differently for different simulations. Anchors are not taken into consideration at this stage, since various simulations demand different anchor placement strategies.

4.2. Upper and lower bound networks

We first define upper and lower bound networks.

Definition 2. Suppose $\mathcal{N}=(A,S,M)$ is a connected network. Given a subset $\tilde{S}\subseteq S$,

1. a upper bound network $\mathcal{N}^{(u)}$ is defined to be $(A^{(u)},\tilde{S},M^{(u)})$, where $A^{(u)}=\{j|\exists\overline{d_{ij}}\in M$, where $i\in\tilde{S},j\in A\}$ and $M^{(u)}=\{\overline{d_{ij}}|\exists\overline{d_{ij}}\in M$, where $i,j\in\tilde{S}\cup A^{(u)}\}$;
2. a lower bound network $\mathcal{N}^{(l)}$ is defined to be $(A^{(l)},\tilde{S},M^{(l)})$, where $A^{(l)}=A^{(u)}\cup\{i|(\exists\overline{d_{ij}}\in M)\vee(\exists\overline{d_{ji}}\in M)$, where $i\in(S\backslash\tilde{S}),j\in\tilde{S}\}$ and $M^{(l)}=\{\overline{d_{ij}}|\exists\overline{d_{ij}}\in M$, where $i,j\in\tilde{S}\cup A^{(l)}\}$;

We define a *relevant* distance measurement as a distance measurement in M involving a node in $\tilde{S}$ and a *relevant* node as a node involved in a *relevant* distance measurement. Then, for the upper bound network $\mathcal{N}^{(u)}$, $A^{(u)}$ consists of all *relevant* anchors and $M^{(u)}$ consists of all *relevant* distance measurements between any two nodes in $\tilde{S}$ and $A^{(u)}$; for the lower bound network $\mathcal{N}^{(l)}$, $A^{(l)}$ consists of all *relevant* nodes excluded from $\tilde{S}$ and $M^{(l)}$ consists of all the *relevant* distance measurements. Fig. 3 shows examples of an upper and a lower bound networks associated with a sensor subset $\tilde{S}$. Clearly, both of them are determined by local topologies involving $\tilde{S}$; if the upper bound network is connected, the corresponding lower bound network must be connected as well. The following theorem describes the relationships between the CRLBs for sensors in $\tilde{S}$ in the original network and two bound networks.

Theorem 6. *Given a connected network $\mathcal{N}=(A,S,M)$ and a subset $\tilde{S}\subseteq S$, construct an upper bound network $\mathcal{N}^{(u)}=(A^{(u)},\tilde{S},M^{(u)})$ and a lower bound network $\mathcal{N}^{(l)}=(A^{(l)},\tilde{S},M^{(l)})$. If $\mathcal{N}^{(u)}$ is connected and $A^{(u)}\neq\emptyset$, for any sensor $s\in\tilde{S}$, let c_s, $c_s^{(u)}$ and $c_s^{(l)}$ be the CRLBs in the three networks, respectively, and $c_s^{(l)}\le c_s\le c_s^{(u)}$.*

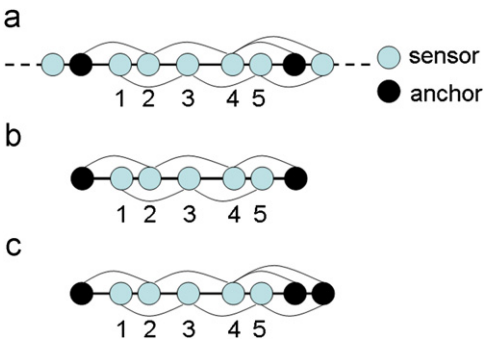


Fig. 3. Examples of an upper and a lower bound networks. Given an original network in (a) and a sensor subset $\tilde{S}=\{1,2,3,4,5\}$, an upper and a lower bound networks are constructed in (b) and (c), respectively.

Proof. Without loss of generality, let $\tilde{S}=\{1,\ldots,p\}$. Define the FIMs of the three networks as $J_n J_p^{(u)}$ and $J_p^{(l)}$ respectively. Because both $\mathcal{N}$ and $\mathcal{N}^{(u)}$ are connected and $A^{(u)}\neq\emptyset$, the three FIMs are all positive-definite. Then define C_n,$C_p^{(u)}$ and $C_p^{(l)}$ to be their inverses respectively and C_p to be the leading $p\times p$ submatrix of C_n.

In $\mathcal{N}$, replace q sensors not in $\tilde{S}$ but directly connected to sensors in $\tilde{S}$ by additional anchors and the new FIM is a block-diagonal matrix:

$$\tilde{J}_{n-q}=\begin{pmatrix} J_p^{(l)} & \mathbf{0} \\ \mathbf{0}^T & \cdots \end{pmatrix}. \tag{25}$$

Because such replacement neither empties the anchor set nor disconnects the network, $\tilde{J}_{n-q}$ is positive-definite:

$$\tilde{J}_{n-q}^{-1}=\begin{pmatrix} J_p^{(l)-1} & \mathbf{0} \\ \mathbf{0}^T & \cdots \end{pmatrix}. \tag{26}$$

From Theorem 5, every diagonal entry in $C_p^{(l)}$ is less than or equal to the corresponding entry in C_n. Thus, $c_s^{(l)}\le c_s$. Secondly, we can transform $\mathcal{N}^{(u)}$ into $\mathcal{N}$ by adding sensors, anchors and distance measurements. According to Theorems 2–4, all these operations do not increase the CRLBs for sensors in S_p; thus $c_s\le c_s^{(u)}$. □

Remark 9. Theorem 6 indicates that an upper and a lower bounds on the CRLB for a sensor can be derived provided that the topology information within a certain limited range around this sensor is known. Consider a sensor s in a network $\mathcal{N}=\{A,S,M\}$ and two sensor subsets $\tilde{S}$ and $\overline{S}$ satisfying $s\in\tilde{S}$ and $\tilde{S}\subset\overline{S}\subset S$. Let $\mathcal{N}^{(u)}$ and $\overline{\mathcal{N}}^{(u)}$ be upper bound networks associated with the two subsets. Suppose that each of them is connected and contains at least one anchor. Let c_s be the CRLB for sensor s in $\mathcal{N}$ and $c_s^{(u)}$ and $\overline{c_s}^{(u)}$ be the CRLBs for sensor s in $\mathcal{N}^{(u)}$ and $\overline{\mathcal{N}}^{(u)}$ respectively. According to Theorem 6, both $c_s^{(u)}$ and $\overline{c_s}^{(u)}$ are upper bounds on c_s. Because of $\tilde{S}\subset\overline{S}$, $\mathcal{N}^{(u)}$ can be transformed into $\overline{\mathcal{N}}^{(u)}$ by adding sensors, anchors and associated distance measurements; according to Theorems 2–4, we can obtain $c_s^{(u)}\ge\overline{c_s}^{(u)}\ge c_s$, implying that the upper bound based on a large subset is tighter

than that based on a small subset. The same conclusion holds for the lower bound case.

4.3. Partitioning networks

We can employ anchors as boundaries to partition a connected network $\mathcal{N}$ into a group of segments such that sensors belonging to such a segment form a sensor subset. The union of all subsets equals the sensor set of $\mathcal{N}$ and the intersection of any two subsets is empty. Given a sensor subset in $\mathcal{N}$, we establish an upper bound network $\mathcal{N}^{(u)}$ and a lower bound network $\mathcal{N}^{(l)}$, which are both unit networks; the CRLB for each sensor in the subset can be bounded using these two bound networks. In fact, we can transform $\mathcal{N}^{(u)}$ to $\mathcal{N}^{(l)}$ by placing (maybe zero) additional anchors on the boundaries of $\mathcal{N}^{(u)}$ with anchors nearby and adding distance measurements associated with these anchors. According to Remark 7, additional anchors near existing anchors have little effects on reducing the CRLBs and consequently, the CRLBs in $\mathcal{N}^{(u)}$ are close to those in $\mathcal{N}^{(l)}$ such that both bounds are tight. *Because the tight bounds for a sensor are derived based on the information associated with the sensor subset including this sensor, we are ensured that the localization error of this sensor is mainly determined by the local part of the whole network.*

Simulations are conducted to verify the effectiveness of this partitioning strategy. In each generated valid network, we place m equally spaced anchors in the range of $[0, L]$ with two of them at 0 and L respectively. Then, such a valid network is partitioned into $m-1$ segments by m anchors; for each segment, we can build a pair of upper and lower bound networks. Given $L=100$ and $m=10$, we take into account different values of n in the simulations. As can be seen from Fig. 4, the CRLB for every sensor in a valid network is tightly bounded by the CRLBs in the corresponding bound networks.

In essence, this partitioning strategy amounts to a divide and conquer method in that the analysis of localization errors in a large-scale network can be converted into the analysis in small-scale unit networks. If these unit networks have network parameters in common, we only need to analyze one of them, instead of all of them.

4.4. Localization errors in unit networks

At first, a class of regular unit networks are shown in Fig. 5, and the regularity means: (1) *each sensor receives h distance measurements from its left neighboring nodes and another h distance measurements from its right neighboring nodes*; (2) *to guarantee the sensors close to boundaries can also receive h distance measurements from each side, h anchors have to be placed at the left side of all sensors and another h anchors at the right side*.

It turns out that sensors have the identical *node degree*, i.e. 2 h, in a regular unit network. In a regular unit network with node degree 2, 4, 6, $2n$ (n equals the number of sensors), if the standard deviations of distance measurement errors are equal, closed-form formulas for the CRLBs presented in [29] exactly characterize the behaviors of localization errors. Furthermore, the following theorem helps us understand error propagation in general unit networks based on regular unit networks.

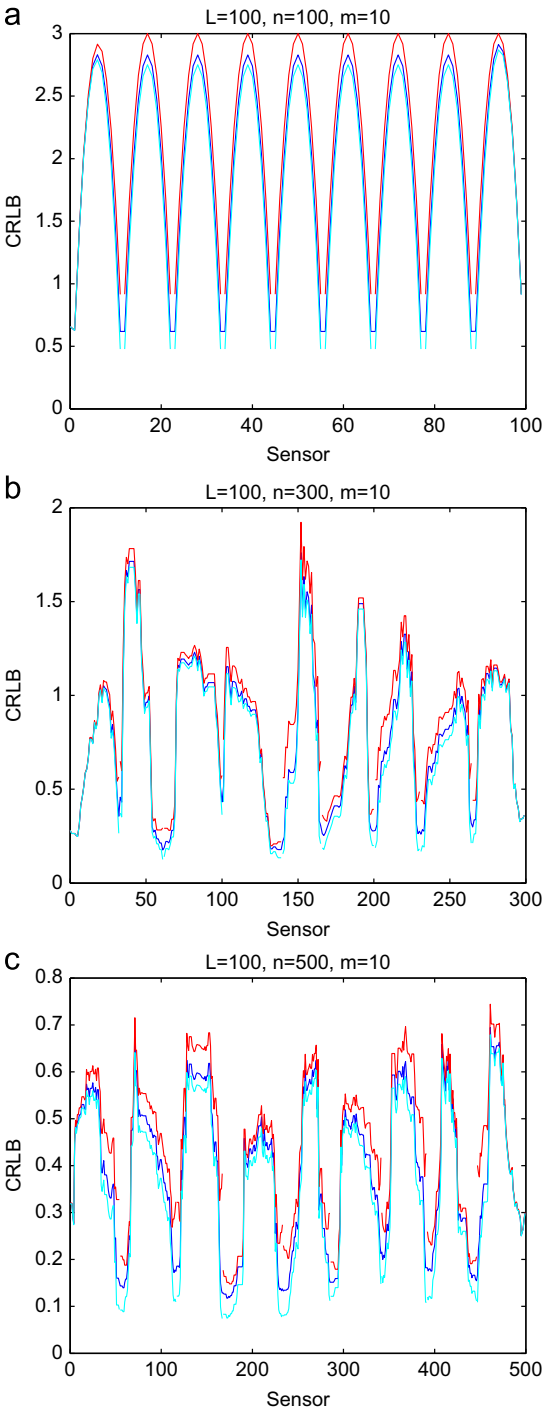


Fig. 4. The CRLBs in actual networks (blue, middle), upper bound networks (red, top) and lower bound networks (cyan, bottom) with respect to different sensor densities. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

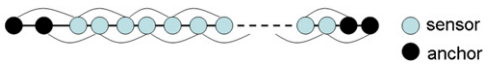


Fig. 5. A regular unit network with $h=2$.

Theorem 7. *$\mathcal{N}=(A,S,M)$ $(A\neq\emptyset)$ is a connected unit network. Let d_{max} and d_{min} be the maximum and minimum node degrees of the graph corresponding to $\mathcal{N}$. Construct two regular unit networks $\mathcal{N}^{(k_1)}$ and $\mathcal{N}^{(k_2)}$ with the same number of sensors as in $\mathcal{N}$ and with unique node degrees k_1 and k_2 $(k_1\le k_2)$, respectively. Let C_n, $C_n^{(k_1)}$ and $C_n^{(k_2)}$ be the inverses of the FIMs for the three networks. If $k_1\le d_{min}\le d_{max}\le k_2$, then for every sensor s $(1\le s\le n)$ $(C_n^{(k_1)})_{ss}\ge(C_n)_{ss}\ge(C_n^{(k_2)})_{ss}$.*

Proof. If $k_1\le d_{min}\le d_{max}\le k_2$, $\mathcal{N}^{(k_1)}$ can be transformed into $\mathcal{N}$ by adding anchors and distance measurements between sensors or between sensors and anchors and based on Theorems 2 and 4, these modifications do not decrease the CRLBs and thus $(C_n^{(k_1)})_{ss}\ge(C_n)_{ss}$. Similarly, we have $(C_n)_{ss}\ge(C_n^{(k_2)})_{ss}$. □

Remark 10. Theorem 7 indicates that in a unit network, the CRLB for each sensor can be bounded using regular unit networks. If $d_{max}-d_{min}$ is small, so is k_2-k_1; the bounds are tight and thus helpful. But it is less probable for $d_{max}-d_{min}$ to be small in a unit network in which sensors are randomly and uniformly distributed. Supposing the mean and standard deviation of node degrees in a unit network to be x and δ, one could postulate that an approximation to performance would follow from a regular unit network calculation in which the parameters d_{min} and d_{max} are approximated by $\overline{d_{min}}=x-\delta$ and $\overline{d_{max}}=x+\delta$ respectively.

Given $L=100$ and different values of n, simulations are conducted to validate this approximation method by generating four valid unit networks with two anchors at 0 and L respectively. As shown in Fig. 6, $\overline{d_{min}}$ and $\overline{d_{max}}$ produce tighter bounds than d_{min} and d_{max}. Although the curves corresponding to the CRLBs in the generated valid unit networks are not as smoothly parabolic as in regular unit networks, they share *two common properties that the CRLBs increase at a quadratic speed and the sensors with the maximum CRLBs lie around the middle of unit networks*. Moreover, the upper and lower bounds based on regular unit networks allow us to predict how fast localization errors grow in a unit network given the knowledge about sensor density which determines both x and δ.

5. Further discussion

In this section, two practical issues are investigated.

5.1. Applicability of the analysis based on the CRLB

An analysis based on the CRLB is often faced with a critical question whether the CRLB is obtainable in actual localization procedures, which determines the applicability of the analysis. We first consider a localization procedure in which all sensors are localized simultaneously using the MLE. Given a one-dimensional network $\mathcal{N}=(A,S,M)$, x_i and x_j are true locations of i and j; if $x_i>x_j$, the distance measurement $\overline{d_{ij}}$ can be expressed as follows:

$$\overline{d_{ij}}=x_i-x_j+e_{ij}. \tag{27}$$

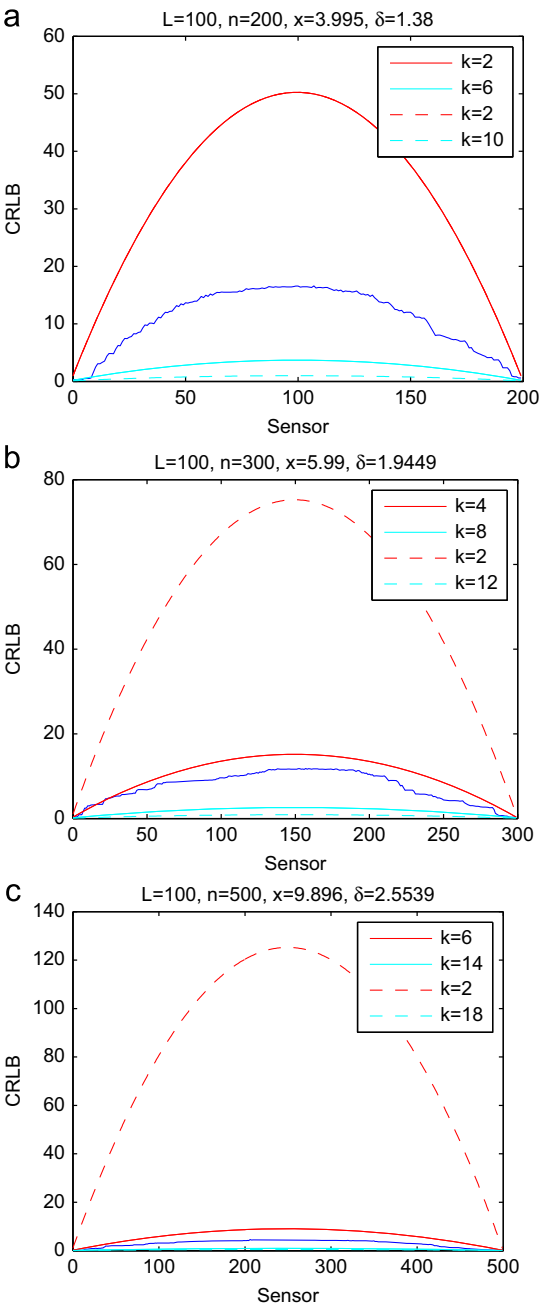


Fig. 6. The actual CRLBs and their upper bounds and lower bounds in unit networks. The actual CRLBs are plotted in solid blue curves. Upper bounds with red dashed curves are produced using d_{min}, lower bounds with cyan dashed curves using d_{max}, upper bounds with red solid curves using $\overline{d_{min}}$ and lower bounds with cyan solid curves using $\overline{d_{max}}$. Note that the red solid and the red dashed curves overlap in (a). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Considering distance measurements from at least one sensor, the corresponding equations like (27) can be stacked together and written in vector form

$$\mathbf{z}=(P\ \ Q)\begin{pmatrix}\mathbf{x}\\ \mathbf{a}\end{pmatrix}+\mathbf{e}, \tag{28}$$

where $\mathbf{z}$ is the distance measurement vector of order s, say; $\mathbf{a}$ is the anchor location vector of order m; P and Q are $s\times n$ and $s\times m$ matrices consisting of only 0, -1, 1; $\mathbf{e}$ is the error vector associated with $\mathbf{z}$; $\mathbf{x}$ is the vector containing n parameters (i.e. sensor locations) to be estimated. Equivalently, (28) can be expressed as

$$\mathbf{z}-Q\mathbf{a}=P\mathbf{x}+\mathbf{e}. \tag{29}$$

Then, the generalized least squares method finds estimate $\hat{\mathbf{x}}$ by solving

$$\hat{\mathbf{x}}=\arg\min_{\mathbf{x}}[(\mathbf{z}-Q\mathbf{a}-P\mathbf{x})^T(Var(\mathbf{e}))^{-1}(\mathbf{z}-Q\mathbf{a}-P\mathbf{x})]. \tag{30}$$

Because $\mathbf{e}$ is Gaussian, the same estimate $\hat{\mathbf{x}}$ can also be derived by the MLE. Moreover, the estimate $\hat{\mathbf{x}}$ has an explicit formula

$$\hat{\mathbf{x}}=(P^T(Var(\mathbf{e}))^{-1}P)^{-1}P^T(Var(\mathbf{e}))^{-1}(\mathbf{z}-Q\mathbf{a}), \tag{31}$$

and its covariance matrix is

$$Var(\hat{\mathbf{x}})=(P^T(Var(\mathbf{e}))^{-1}P)^{-1}. \tag{32}$$

Regarding P, its row number s equals the number of distance measurements from at least one sensor and its column number n equals the number of sensors. For the k-th distance measurement, if it involves two sensors i and j, $(P)_{ki}=1$, $(P)_{kj}=-1$ (or $(P)_{ki}=-1$ and $(P)_{kj}=1$ which does not change the matrix $P^T(Var(\mathbf{e}))^{-1}P$) and all the other entries in this row are 0; if it involves one sensor i and one anchor j, $(P)_{ki}=1$ (or $(P)_{ki}=-1$ which neither changes the matrix $P^T\,Var(\mathbf{e})P$) and all the other entries in this row are 0. Moreover, due to the assumption of independent errors, $Var(\mathbf{e})$ is an $s\times s$ diagonal matrix. Then, it is straightforward to verify that $P^T(Var(\mathbf{e}))^{-1}P$ equals J_n. Thus, the MLE attains the CRLB.

Therefore, our analysis is certainly applicable here. But the above localization procedure essentially depends on centralized processing which is impractical for large-scale networks. Decentralized processing can in fact be considered also.

Recalling the construction of upper bound networks in a large-scale network introduced in Section 4.3, it is clear that each upper bound network is actually a subnetwork (or subset) of the original large-scale network. Assuming that one node is elected from such a subnetwork and collects all information within this subnetwork, including available distance measurements and anchor positions, a locally centralized localization procedure using the MLE can be performed on this node. Firstly, sensors in this subnetwork are localizable since the corresponding upper bound network is a unit network and consists of two anchors at two ends. Secondly, considering the comparatively small scale of such subnetworks, it is convenient to fulfill the assumption about the node election and information collection such that a decentralized localization procedure is realized. Thirdly, according to the result about the tightness of upper bounds derived from upper bound networks in Section 4.3 which is well reflected in simulations, the localization errors produced in the decentralized localization procedure (i.e. the CRLB in upper bound networks) is much closer to the errors produced in the centralized localization procedure (i.e. the CRLB in the original large-scale network). Hence, the analysis based on the CRLB is also indicative for decentralized processing.

5.2. The optimal placement of anchors

As we mentioned previously, anchor has significant impacts on localization errors. Assuming anchors' locations are controllable, an optimal anchor placement should minimize the total (or average) CRLB in a network, which is intrinsically an optimization problem. Here we give a preliminary discussion based on the preceding analysis.

According to Remark 7, an additional anchor should be placed near the sensor with the largest CRLB in a network to efficiently reduce the total CRLB. Imagining sensors are randomly and uniformly distributed in a network with two anchors at both ends, in the light of the discussion in Section 4.4, the sensor with the largest CRLB will probably lie around the middle of the network, and hence, it is reasonable to place the additional anchor at the middle of the network; by iteratively applying this strategy, anchors are finally equally spaced in the network. Thus, we conjecture that, from the statistical point of view, distributing anchors with equal spacing should be optimal in networks with randomly and uniformly distributed sensors.

In the simulations, we generate 100 valid networks and consider 1000 different anchor placement cases in which apart from the proposed case that anchors are equally spaced, anchors are randomly and uniformly placed. The variance of anchor intervals is defined to be the variance of the lengths between any pair of adjacent anchors, and reflects how far one anchor placement case is away from the equal spacing case; in particular, the variance of anchor intervals for the equal spacing case is 0. Then, we implement each anchor placement case in the localization procedures of the 100 valid networks, and compute the average CRLB. In the end, the simulation results are plotted in Fig. 7. As can be seen, the minimum average CRLB appears to increase almost linearly with the variance of anchor intervals, and the minimum occurs when the variance is 0, which verifies that the equal spacing case is optimal in a statistical sense.

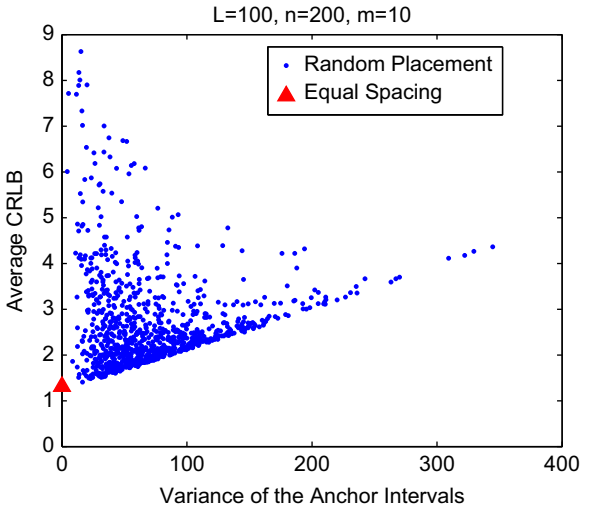


Fig. 7. The average CRLB with different anchor placement strategies.

6. Conclusions

In this paper, we investigated localization errors in one-dimensional sensor networks in terms of the CRLB. We theoretically analyzed how different factors affect the CRLB and thus localization errors, and concluded that the localization error for a sensor is mainly determined by the local elements within a certain limited range around this sensor. The analysis of localization errors in a large-scale network can be broken down into the analysis in a number of small-scale unit networks. Also, we discussed two practical issues, viz. the applicability of the analysis based on the CRLB and the optimal anchor placement.

The physical embodiment of a one-dimensional sensor network does not necessarily involve an ideal straight line. It might involve a large circle, a curving road, an irregular boundary of a region or a coastline for example, as long as the curvature is small. Given a one-dimensional sensor network, the localization task aims to determine the distance from each sensor to some starting point of this network. Provided that this network is deployed over a non-straight line, this distance is then interpreted as the accumulated distance from this sensor to the starting point along this non-straight line. In this case, if the curvature of this line is relatively small and Euclidean distance measurements between any pair of nodes are close to the corresponding accumulated distances, localization results are then consistent with the above interpretation.

Furthermore, although all these results are limited to one-dimensional sensor networks and cannot be easily extended to two- and three-dimensional cases, the treatments we employed here might be still useful. Our analysis relies on the closed-form formulas for the CRLB in one-dimensional regular unit networks, in which FIMs are banded symmetric Toeplitz matrices; in two-dimensional lattice networks with anchors on the boundary and surrounding sensors, FIMs are actually symmetric block-tridiagonal Toeplitz matrices, which makes it possible for us to deal with localization errors in two-dimensional sensor networks in a like manner to that we used in the one-dimensional case.

Acknowledgments

B. Huang, C. Yu and B.D.O. Anderson are supported by the Australian Research Council under DP-110100538. C. Yu is an ARC Queen Elizabeth II Fellow. B.D.O. Anderson and B. Huang are also supported by National ICT Australia-NICTA. NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. This material is based on research sponsored by the Air Force Research Laboratory, under agreement number FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

References

[1] Y. Chen, J. Hwang, A power-line-based sensor network for proactive electrical fire precaution and early discovery, IEEE Transactions on Power Delivery 23 (2) (2008) 633–639.
[2] K. Han, H. Xu, Research on wireless network-based power line inspection, International Forum on Information Technology and Applications 1 (2009) 379–381.
[3] G. Shi, Z. Zhao, Y. Shu, S. Han, Monitoring the high-voltage transmission lines based on two-tier wireless networks, in: Proceedings of the 4th ACM International Workshop on Experimental Evaluation and Characterization, New York, NY, USA, 2009, pp. 77–78.
[4] I. Stoianov, L. Nachman, S. Madden, T. Tokmouline, M. Csail, Pipenet: a wireless sensor network for pipeline monitoring, in: The 6th International Conference on Information Processing in Sensor Networks (IPSN), Cambridge, MA, USA, 2007, pp. 264–273.
[5] S. Yoon, W. Ye, J. Heidemann, B. Littlefield, C. Shahabi, Swats: wireless sensor networks for steamflood and waterflood pipeline monitoring, IEEE Transactions on Signal Processing 55 (2) (2007) 684–696.
[6] I. Jawhar, N. Mohamed, K. Shuaib, A framework for pipeline infrastructure monitoring using wireless sensor networks, in: Wireless Telecommunications Symposium, 2007, pp. 1–7.
[7] A. Ahrary, Localization of a mobile robot in sewer pipes using hough transform, ICGST International Journal on Graphics, Vision and Image Processing, GVIP 9 (2009) 17–23.
[8] X. Niu, X. Huang, Z. Zhao, Y. Zhang, C. Huang, L. Cui, The design and evaluation of a wireless sensor network for mine safety monitoring, in: IEEE Global Telecommunications Conference (GLOBECOM), Washington, DC, USA, 2007, pp. 1291–1295.
[9] P. Li, X. Huang, Y. Fang, P. Lin, Optimal placement of gateways in vehicular networks, IEEE Transactions on Vehicular Technology 56 (6) (2007) 3421–3430.
[10] Wireless wildfire sensor networks, 〈http://www.bushfirecrc.com/research/a51/publicdocuments.html〉.
[11] J. Albowicz, A. Chen, L. Zhang, Recursive position estimation in sensor networks, in: The 9th International Conference on Network Protocols, Riverside, CA, USA, 2001, pp. 35–41.
[12] D. Niculescu, B. Nath, Localized positioning in ad hoc networks, Ad Hoc Networks 1 (2–3) (2003) 247–259 (Special issue on Sensor Network Protocols and Applications).
[13] A. Savvides, H. Park, M. Srivastava, The n-hop multilateration primitive for node localization problems, Mobile Networks and Applications 8 (4) (2003) 443–451.
[14] J. Fang, M. Cao, A.S. Morse, B.D.O. Anderson, Localization of sensor networks using sweeps, in: The 45th IEEE Conference on Decision and Control (CDC), San Diego, CA ,USA, 2006, pp. 4645–4650.
[15] A. Savvides, W. Garber, S. Adlakha, R. Moses, M.B. Srivastava, On the error characteristics of multihop node localization in ad-hoc sensor networks, in: The 2nd International Conference on Information Processing in Sensor Networks (IPSN), vol. 2634, Palo Alto, CA, USA, 2003, pp. 317–332.
[16] A. Savvides, W.L. Garber, An analysis of error inducing parameters in multihop sensor node localization, IEEE Transactions on Mobile Computing 4 (6) (2005) 567–577.
[17] K. Whitehouse, C. Karlof, A. Woo, F. Jiang, D. Culler, The effects of ranging noise on multihop localization: an empirical study, in: The 4th International Conference on Information Processing in Sensor Networks (IPSN), Piscataway, NJ, USA, 2005, pp. 73–80.
[18] G. Frenkel, Geometric dilution of position (GDOP) in position determination through radio signals, Proceedings of the IEEE 61 (4) (1973) 496–497.
[19] I. Sharp, K. Yu, Y. Guo, GDOP analysis for positioning system design, IEEE Transactions on Vehicular Technology 58 (7) (2009) 3371–3382.
[20] C. Chang, A. Sahai, Estimation bounds for localization, in: Proceedings of IEEE Sensor and Ad Hoc Communications and Networks (SECON), Santa Clara, CA, USA, 2004, pp. 415–424.
[21] N. Patwari, A.O. Hero, M. Perkins, N.S. Correal, R.J. O'Dea, Relative location estimation in wireless sensor networks, IEEE Transactions on Signal Processing 51 (8) (2003) 2137–2148.
[22] E.G. Larsson, Cramer–Rao bound analysis of distributed positioning in sensor networks, IEEE Signal Processing Letters 11 (3) (2004) 334–337.
[23] H. Wang, L. Yip, K. Yao, D. Estrin, Lower bounds of localization uncertainty in sensor networks, in: Proceedings of IEEE Acoustics,

Speech, and Signal Processing (ICASSP), Salt Lake City, UT, USA, 2004, pp. 917–920.
[24] D. Niculescu, B. Nath, Error characteristics of ad hoc positioning systems (APS), in: The 5th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc), New York, NY, USA, 2004, pp. 20–30.
[25] Q. Shi, S. Kyperountas, N.S. Correal, F. Niu, Performance analysis of relative location estimation for multihop wireless sensor networks, IEEE Journal on Selected Areas in Communications 23 (4) (2005) 830–838.
[26] R. Huang, G.V. Zaruba, M. Huber, Complexity and error propagation of localization using interferometric ranging, in: IEEE International Conference on Communications (ICC), Glasgow, UK, 2007, pp. 3063–3069.
[27] F. Chang, Inversion of a perturbed matrix, Applied Mathematics Letters 19 (2) (2006) 169–173.
[28] G.H. Golub, C.F.V. Loan, Matrix Computations, third ed., Johns Hopkins University Press, Baltimore, MD, USA, 1996.
[29] B. Huang, C. Yu, B.D.O. Anderson, Error propagation in network localization with regular topologies, in: IEEE Global Telecommunications Conference (GLOBECOM), Honolulu, HI, USA, 2009, pp. 1–6.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2010 proceedings.

Connectivity-based Distance Estimation in Wireless Sensor Networks

Baoqi Huang, Changbin Yu, Brian D.O. Anderson, Guoqiang Mao

***Abstract*—Distance estimation is of great importance for localization and a variety of applications in wireless sensor networks. In this paper, we develop a simple and efficient method for estimating distances between any pairs of neighboring nodes in static wireless sensor networks based on their local connectivity information, namely the numbers of their common one-hop neighbors and non-common one-hop neighbors. The proposed method involves two steps: estimating an intermediate parameter through a Maximum-Likelihood Estimator (MLE) and then mapping this estimate to the associated distance estimate. In the first instance, we present the method by assuming that signal transmission satisfies the ideal unit disk model but then we expand it to the more realistic log-normal shadowing model. Finally, simulation results show that localization algorithms using the distance estimates produced by this method can deliver superior performances in most cases in comparison with the corresponding connectivity-based localization algorithms.**

I. Introduction

Wireless sensor networks, comprised of hundreds or thousands of small and inexpensive nodes with constrained computing power, limited memory and short battery lifetime, can be used to monitor and collect data in a region of interest. Accurate and low-cost sensor (the word "sensor" connotes a node of unspecified location) localization is a critical requirement for a wide variety of applications in wireless sensor networks, and great efforts have been invested in developing localization algorithms including both distance-based algorithms and connectivity-based algorithms.

In reality, exact distance measurement is usually unavailable and has to be estimated from information such as received signal strength (RSS), time of arrival (TOA), or time difference of arrival (TDOA) [1]. In large-scale sensor networks, it is, however, impractical to localize all sensors by using additional hardware such as GPS receivers and measuring devices due to cost constraints. On the other hand, although classes of connectivity-based localization algorithms without using any additional measuring devices have been proposed [2], [3], achieving a high localization accuracy usually demands a comparatively large number of anchor nodes, hereafter termed simply anchors, whose positions are known a priori.

In a static wireless sensor network, two nodes are termed one-hop neighbors or simply neighbors as long as they can communicate with each other. An intuitive observation shows that two geographically close neighboring nodes often share more common one-hop neighbors than two distant nodes. In this paper, we quantify and exploit this observation to develop a method for estimating the distance between any pair of neighboring nodes based on their local connectivity information, i.e. the numbers of their common one-hop neighbors and non-common one-hop neighbors. Our method involves two steps: first, an intermediate parameter relating the distance and the numbers of different types of neighbors is estimated based on a Maximum Likelihood Estimator (MLE); second, through a mapping function we can obtain the distance estimate from the estimate in the first step. After presenting this method for the unit disk model, we expand it to the more realistic log-normal shadowing model. Such distance estimates can be directly used by distance-based localization algorithms, and in comparison with traditional connectivity-based localization algorithms, a significant improvement on localization accuracies is reported through simulations.

The advantages of the proposed method are: independent of additional hardware; totally distributed; energy efficient due to its simple mechanism and computations. Prior to our work, [4], [5] came up with the method of estimating distances based on the same idea as ours; their treatments rest on empirical observations rather than theoretical foundations. In comparison to their work, our paper: (1) formalizes and gives mathematical proofs of the correctness of the method; (2) bases the method on the MLE in both the unit disk model and the more realistic log-normal shadowing model; (3) reports the performance improvement on localization by using the distance estimates produced by the method through simulations.

The remainder of the paper is organized as follows: Section II introduces the method of estimating distances in the unit disk model and Section III expands it to the log-normal shadowing model. Section IV investigates the performance improvement on localization by using the proposed method through simulations. Finally, Section V concludes the paper.

II. Estimating Distances in the Unit Disk Model

We first introduce the network model used here and then present the method in the unit disk model.

A. Network Model

Considering a static wireless sensor network where nodes are randomly and uniformly distributed in 2-dimensional region, a homogeneous Poisson process provides an accurate model for the distribution of nodes as the network size approaches infinity [6]. Let λ denote the node density (the number of nodes per unit square) and the probability mass function of the number of nodes N in an area D is given by

$$Pr(N = n) = \frac{(\lambda D)^n}{n!} e^{-\lambda D} \tag{1}$$

B. Huang and B.D.O. Anderson are with the Australian National University and National ICT Australia Ltd., Canberra. Email:{Baoqi.Huang,Brian.Anderson}@anu.edu.au.

C. Yu is with the Australian National University, Canberra. Email:Brad.Yu@anu.edu.au. Correspondence: RSISE Building 115, The Australian National University, Canberra ACT 0200, Australia, Tel: +61 2 61258670, Fax: +61 2 61258660.

Guoqiang Mao is with the University of Sydney and National ICT Australia Ltd., Sydney. Email:Guoqiang@ee.usyd.edu.au.

978-1-4244-5638-3/10/$26.00 ©2010 IEEE

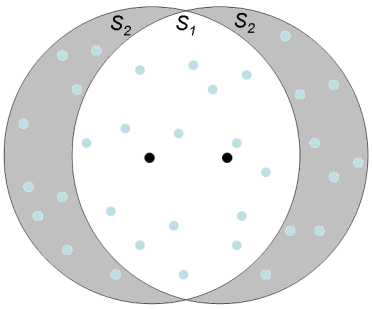


Fig. 1. Communication coverage of two neighboring nodes in the unit disk model.

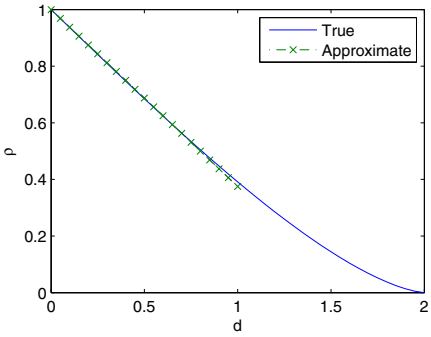


Fig. 2. The true and approximate functions from d to ρ ($r = 1$).

where $Pr(\cdot)$ denotes the probability of a statistical event. Obviously, N is a Poisson random variable with mean λD.

Moreover, throughout this paper we assume that:

Assumption 1: Nodes are Poissonly distributed with a uniform density λ in an infinite plane.

Assumption 2: Each node knows the neighborhood information (i.e. the list of one-hop neighbors) of all its one-hop neighbors.

Assumption 1 avoids the boundary effect and Assumption 2 ensures that local connectivity information is available for each node to carry out the distance estimation method.

B. Estimating Distances

Given a static wireless sensor network conforming Assumption 1, the node transmission range is identically r as we are discussing with the unit disk model.

As shown in Fig. 1, provided two neighboring nodes A and B separated by distance d ($d \leq r$), two circles with radii r and centered at A, B represent their individual communication coverage, and intersect and create three disjoint regions. The nodes residing in the middle one are common one-hop neighbors of A and B and the nodes residing in the left (right) one are non-common one-hop neighbors of A (B). Define S_1 to be the area of the middle one and both the areas of the left and the right ones are $\pi r^2 - S_1$, denoted S_2. Moreover, we let three random variables M, P, Q denote the numbers of the three kinds of neighbors and define a key parameter ρ:

$$\rho = \frac{E(2M)}{E(2M+P+Q)} \tag{2}$$

where $E(\cdot)$ denotes the expected value of a random variable.

As pointed out in [7], M, P, Q are mutually independent Poisson distributed random variables with the means $\lambda S_1, \lambda S_2, \lambda S_2$. Equivalently, ρ can also be expressed as

$$\rho = \frac{S_1}{S_1 + S_2} = \frac{S_1}{\pi r^2} \tag{3}$$

According to the geometries among d, r and S_1, we have

$$S_1 = 2r^2 \arccos(\frac{d}{2r}) - d\sqrt{r^2 - \frac{d^2}{4}} \tag{4}$$

In effect, S_1 is a monotonically decreasing function of d, and so is ρ. Hence, as long as ρ is available, we can compute d by using its inverse function, termed a mapping function.

The actual values of M, P and Q can be easily obtained in the light of Assumption 2 and can be furthered employed to estimate ρ according to Theorem 1.

Theorem 1: Given three independent Poisson distributed random variables M, P and Q which define $\rho = \frac{E(2M)}{E(2M+P+Q)}$, the MLE for ρ, termed $\hat{\rho}$, is

$$\hat{\rho} = \begin{cases} 1, & \text{if M=P=Q=0} \quad (5) \\ \dfrac{2M}{2M+P+Q}, & \text{otherwise} \quad (6) \end{cases}$$

Proof: Establish a statistical model: measured data are observations of M, P and Q, denoted $\phi = [\ m \quad p \quad q\ ]$ where m, p, q are non-negative integers; the unknown parameters are $\theta = [\ \rho \quad \lambda\ ]$. The likelihood function of this model is

$$\mathcal{L}(\theta, \phi) = Pr(M = m) \times Pr(P = p) \times Pr(Q = q) \tag{7}$$

The MLE is the solution to the following equation set

$$\frac{\partial \ln \mathcal{L}(\theta, \phi)}{\partial \theta} = 0 \tag{8}$$

which yields

$$\begin{cases} \dfrac{m}{\rho} - \dfrac{p+q}{1-\rho} + \lambda\pi r^2 = 0 & (9) \\ \dfrac{m+p+q}{\lambda\pi r^2} - (2-\rho) = 0 & (10) \end{cases}$$

By eliminating λ, we can obtain

$$2m = (2m + p + q)\rho \tag{11}$$

If $2m + p + q > 0$, i.e. m, p and q are simultaneously 0, the solution for ρ is $\frac{2m}{2m+p+q}$; otherwise, the solution for ρ is not well-defined. But because $\rho = 1$ maximizes the likelihood when $2m+p+q = 0$, we obtain the MLE for ρ, termed $\hat{\rho}$, as

$$\hat{\rho} = \begin{cases} 1, & \text{if M=P=Q=0} \\ \dfrac{2M}{2M+P+Q}, & \text{otherwise} \end{cases}$$

■

By substituting (4) into (3) and applying the first order Taylor series expansions on $d = 0$, we can obtain

$$\rho \approx 1 - \frac{2d}{\pi r} \tag{12}$$

As depicted in Fig. 2, (12) displays a good approximation to the true function from d to ρ when $0 \leq d \leq r$. As such, the mapping function is approximately

$$\hat{d} \approx \frac{\pi r}{2}(1 - \hat{\rho}) \tag{13}$$

which enables us to obtain the estimate of d, i.e. $\hat{d}$, from $\hat{\rho}$.

III. Estimating Distances in the Log-normal Shadowing Model

Before expanding the method to the more realistic log-normal shadowing model, we make the following assumptions (as is commonly the case in the literature):

Assumption 3: The attenuations of the transmit powers between any pairs of nodes are independent and identically distributed (i.i.d.) [1];

Assumption 4: Communication links are symmetric, namely that node v can directly receive packets from node u as long as node u can directly receive packets from node v;

Assumption 5: All nodes transmit at a fixed power level.

Although these assumptions may not fully reflect a real network environment, they still enable us to obtain some results as estimates for more realistic situations.

A. Log-normal Shadowing Model

The *log-normal shadowing model* predicts the received signal power by a receiver with distance d from a transmitter, denoted $P_R(d)$, to be log-normally distributed around the ensemble average received power, denoted $\overline{P_R(d)}$; this model is based on a wide variety of measurement results [11] as well as analytical evidence [12], and is typically modeled as [1]

$$P_R(d)(dBm) = \overline{P_R}(d_0)(dBm) - 10\alpha \log \frac{d}{d_0} + Z \tag{14}$$

where Z is a random variable representing the shadowing effect, normally distributed with mean zero and variance σ^2; $\overline{P_R}(d_0)$ (dBm) is the ensemble average received signal power in dBm at a short reference distance d_0; α is the path-loss exponent. Both σ and α are a priori known constants; typically, σ is as low as 4 and as high as 12, and α varies between 2 in free space to 6 in heavily built urban areas [11].

Given a transmitter A and a receiver B, if $P_R(d)$ is no less than some specified value P_c, a directional communication link exists from A to B; equivalently, a directional communication link also exists from B to A due to Assumption 4. In particular, if the shadowing effect Z vanishes, i.e. $\sigma = 0$, the log-normal shadowing model is equivalent to the unit disk model with the communication range of

$$r = d_0(\frac{\overline{P_R}(d_0)}{P_c})^{\frac{1}{\alpha}} \tag{15}$$

which is a constant given α. Otherwise, the probability that two nodes with distance d can communicate is

$$g(d) = \int_{k_1 \ln \frac{d}{r}}^{\infty} \frac{e^{-\frac{z^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dz \tag{16}$$

where $k_1 = \frac{10\alpha}{\ln 10}$.

B. Distributions of M, P, Q

We still use M, P, Q to denote the numbers of common and non-common one-hop neighbors associated with A and B. The following theorem and corollary provide their distributions in the log-normal shadowing model.

Theorem 2: Suppose a sensor network where nodes are randomly and uniformly distributed with density λ in a disk of radius R; given two nodes A and B, let M be the number of their common one-hop neighbors and P and Q be the numbers of their non-common one-hop neighbors. M, P and Q are Poisson random variables in the limiting case of $R \to \infty$.

Proof: Let (x_1, y_1) and (x_2, y_2) be the positions of A and B. Given an arbitrary node C, there exist four cases with regard to communications between A, B and C:

1) C can directly communicate with both A and B;
2) C can directly communicate with A but not B;
3) C can directly communicate with B but not A;
4) C cannot directly communicate with either A or B.

Apparently, M is the number of nodes satisfying the Case 1; P (or Q) is the number of nodes satisfying the Case 2 (or 3). Supposing the disk area is D, the probabilities that C satisfies the i-th ($i = 1, 2, 3, 4$) case, termed p_i, are

$$p_1 = \frac{1}{D}\int\int_D g(d_1)g(d_2)dxdy \tag{17}$$

$$p_2 = \frac{1}{D}\int\int_D g(d_1)(1 - g(d_2))dxdy \tag{18}$$

$$p_3 = \frac{1}{D}\int\int_D (1 - g(d_1))g(d_2)dxdy \tag{19}$$

$$p_4 = \frac{1}{D}\int\int_D (1 - g(d_1))(1 - g(d_2))dxdy \tag{20}$$

where $d_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2}$ and $d_2 = \sqrt{(x - x_2)^2 + (y - y_2)^2}$.

To obtain the number of one-hop neighbors of A, i.e. $M+P$, we conduct a test for each node in the network except for A and B to decide whether it satisfies the Case 1 or 2, and then the total number of successful tests is $M + P$. Due to Assumption 3, all of the tests are independent of each other and hence, the test process is a *Bernoulli process*. Then, $M+P$ follows a Binomial distribution with the total number of tests $n = \lambda D - 2$ and the success probability $p = p_1 + p_2$. Moreover, if the limiting case of $\lim_{n\to\infty} np$ converges, the distribution of $M + P$ is Poisson with expected value $\lim_{n\to\infty} np$.

[1] Even though field measurements in real applications seem to indicate that the attenuations between two links with a common node are correlated [8], this i.i.d assumption is generally considered appropriate for far field transmission and is widely used in the literature [8]–[10].

Because of

$$\lim_{n\to\infty} np = \lim_{D\to\infty}[\lambda \int\int_D g(d_1)dxdy] \quad (21)$$

, we let (x_1, y_1) be the origin and $d_1 = R$, and transform (21) into the Polar coordinate system

$$\begin{aligned}\lim_{n\to\infty} np &= \lim_{R\to\infty}[\lambda \int_0^R \int_0^{2\pi} g(x)xdxd\theta] \\ &= 2\pi\lambda \int_0^\infty x \int_{\frac{k_1}{\sigma}\ln\frac{x}{r}}^\infty \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dzdx \quad (22)\end{aligned}$$

When x is no less than some value, say a, $\frac{k_1}{\sigma}\ln\frac{x}{r} \geq 1$ and

$$\begin{aligned}\lim_{n\to\infty} np &= 2\pi\lambda[\int_0^a x \int_{\frac{k_1}{\sigma}\ln\frac{x}{r}}^\infty \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dzdx \\ &\quad + \int_a^\infty x \int_{\frac{k_1}{\sigma}\ln\frac{x}{r}}^\infty \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dzdx] \\ &\leq 2\pi\lambda[\int_0^a x \int_{\frac{k_1}{\sigma}\ln\frac{x}{r}}^\infty \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dzdx \\ &\quad + \int_a^\infty x \int_{\frac{k_1}{\sigma}\ln\frac{x}{r}}^\infty \frac{ze^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dzdx] \quad (23)\end{aligned}$$

It is easy to judge that the two integrals in the right hand side of the inequality definitely converge and then $\lim_{n\to\infty} np < \infty$. Therefore, $M + P$ follows a Poisson distribution. We can obtain the same result for $M+Q$. Regarding M, P, Q, because their success probabilities p_1, p_2 and p_3 in the corresponding Bernoulli processes are less than that of $M+P$, $\lim_{n\to\infty} np < \infty$ when $p = p_1, p_2, p_3$ and thus it is straightforward that M, P and Q follow Poisson distributions. ■

Corollary 1: M, P and Q are mutually independent in the limiting case of $R \to \infty$.

Proof: Define two events $M \leq m$ and $Q \leq q$. Then,

$$\begin{aligned}& Pr(M \leq m \cap Q \leq q) \\ =& \sum_{j=0}^{q} Pr(M \leq m \cap Q = j) \\ =& \sum_{j=0}^{q}[Pr(M \leq m|Q = j)Pr(Q = j)] \\ =& \sum_{j=0}^{q}\sum_{i=0}^{m}[Pr(M = i|Q = j)Pr(Q = j)] \quad (24)\end{aligned}$$

Consider the limiting case of $R \to \infty$ (equivalently $n \to \infty$)

$$\begin{aligned}& \lim_{n\to\infty} Pr(M \leq m \cap Q \leq q) \\ =& \lim_{n\to\infty} \sum_{j=0}^{q}\sum_{i=0}^{m}[Pr(M = i|Q = j)Pr(Q = j)] \\ =& \lim_{n\to\infty} \sum_{j=0}^{q}\sum_{i=0}^{m}[\frac{((n-j)\frac{p_1}{1-p_3})^i e^{-(n-j)\frac{p_1}{1-p_3}}}{i!}\frac{(np_3)^j e^{-np_3}}{j!}] \quad (25)\end{aligned}$$

According to the proof of Theorem 2, when $p = p_1, p_2, p_3$, $\lim_{n\to\infty} np < \infty$, and then $\lim_{n\to\infty} p = 0$. Hence,

$$\begin{aligned}& \lim_{n\to\infty} Pr(M \leq m \cap Q \leq q) \\ =& \lim_{n\to\infty}[\sum_{j=0}^{p}\sum_{i=0}^{m}[\frac{(np_1)^i e^{-np_1}}{i!} \times \frac{(np_3)^j e^{-np_3}}{j!}]] \\ =& \lim_{n\to\infty}[\sum_{i=0}^{m}[\frac{(np_1)^i e^{-np_1}}{i!}] \times \sum_{j=0}^{p}[\frac{(np_3)^j e^{-np_3}}{j!}]] \\ =& \lim_{n\to\infty}[\sum_{i=0}^{m} Pr(M = i) \times \sum_{j=0}^{q} Pr(Q = j)] \\ =& \lim_{n\to\infty}[Pr(M \leq m) \times Pr(Q \leq q)] \quad (26)\end{aligned}$$

Therefore, M and Q are independent as $R \to \infty$. Similarly, we can obtain that M, P and Q are mutually independent through the similar approach. ■

In the log-normal shadowing model, Theorem 2 and Corollary 1 assure us to apply Theorem 1 to arrive at the estimate of ρ. According to [10], the number of one-hop neighbor of a node in the log-normal shadowing model equals $\lambda\pi r^2 e^{\frac{2\sigma^2}{k_1^2}}$, which is in fact $E(M + P)$ and $E(M + Q)$. Furthermore, based on the proof of Theorem 2, we have

$$E(M) = \lambda \int_0^\infty \int_0^{2\pi} g(x)g(\sqrt{x^2 + d^2 - 2xd\cos\theta})xdxd\theta \quad (27)$$

And then we have

$$\rho = \frac{\int_0^\infty \int_0^{2\pi} g(x)g(\sqrt{x^2 + d^2 - 2xd\cos\theta})xdxd\theta}{\pi r^2 e^{\frac{2\sigma^2}{k_1^2}}} \quad (28)$$

which formalizes the functional relationship between ρ and d. Though (28) is not closed-form, we can produce a piecewise linear function to approximate the inverse function of (28) as the mapping function and then estimate d from $\hat{\rho}$.

IV. Simulations

We conduct simulations in Matlab. Since distance estimates are available, it is feasible to localize sensors by using a variety of distance-based localization algorithms. In order to evaluate the proposed method in a fair environment, we shall investigate two well-known connectivity-based localization algorithms which are also applicable with distance measurements, i.e. DV-hop and MDS-MAP; their distance-based versions are termed DV-distance and MDS-MAP distance respectively. Due to the length limitation, refer [2], [3] for more details about these localization algorithms.

To avoid the boundary effect, we actually generate sensor networks over a large square with size of 18×18, but only localize the nodes inside a small square with size of 6×6 and centered at the same center of the large square. However, the nodes outside of the small square are sometimes used in estimating distances between nodes within the small square. Four nodes closest to the four corners of the small square and inside of the small square are chosen as anchors.

In the simulations, the path-loss exponent α is known to be 4, and without loss of generality, the other quantities in (14) describing the log-normal shadowing model, including $\overline{P_R}(d_0)$, d_0 and P_c, are assigned proper values such that r equals 1 according to (15). Equivalently, d and resulting position estimation errors are normalized.

Furthermore, λ takes values of $4, 6, 8, 10$ and σ values of $0, 8$. For each pair of λ and σ, 100 independent runs are carried out and each run involves three steps: generating a wireless sensor network by a homogeneous Poisson process of density λ; estimating distances between any pair of neighboring nodes by using the proposed method; localizing sensors by DV-hop and MDS-MAP and the distance-based versions (with distances coming from the second step), and computing the average position estimation error for each localization algorithm. Finally, the average position estimation errors are plotted, averaged over the 100 independent runs corresponding to each pair of λ and σ and each localization algorithm.

Simulation results plotted in Fig. 3 show that, except the case of $\sigma = 0$ where DV-distance marginally outperforms DV-hop, both DV-distance and MDS-MAP distance using the distance estimates from the proposed method dramatically outperform the corresponding DV-hop and MDS-MAP, and the reduction in the average position estimation errors is at least 30%. Moreover, λ plays an important role in applying distance estimates produced by our method in the localization. In particular, when λ rising from 4 to 6 the average position estimation errors by using DV-distance and MDS-MAP distance are significantly reduced; when λ rises further the average position estimation errors decrease slowly or even increase.

In summary, our distance estimation method makes good use of connectivity in wireless sensor networks and benefits the sensor network localization.

V. Conclusions

In this paper, we proposed a method of estimating distances via connectivity in wireless sensor networks by dealing with the ideal unit disk model and the more realistic log-normal shadowing model. Simulation results showed that using the distance estimates produced by this method significantly improves localization accuracies in comparison with connectivity-based localization algorithms. Future work will focus on analyzing its performance, advancing its practicality by relaxing some assumptions and improving its accuracy by utilizing more information.

Acknowledgements

The authors are supported by the Australian Research Council (ARC) under DP-0877562. C. Yu is an ARC Australian Postdoctoral Fellow. B.D.O. Anderson, G. Mao and B. Huang are also supported by National ICT Australia-NICTA. This material is based on research sponsored by the Air Force Research Laboratory, under agreement number FA2386-09-1-4136.

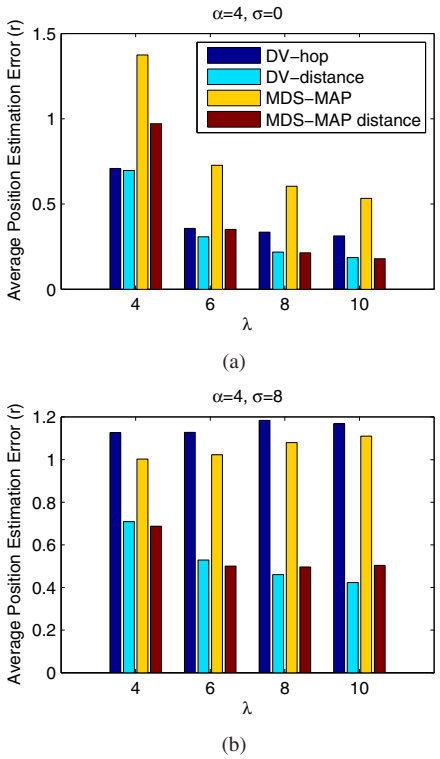


Fig. 3. Average position estimation errors.

References

[1] N. Patwari, J.N. Ash, S. Kyperountas, A.O. Hero III, R.L. Moses, and N.S. Correal. Locating the nodes: cooperative localization in wireless sensor networks. *IEEE Signal Processing Magazine*, 22(4):54–69, Jul. 2005.

[2] D. Niculescu and B. Nath. Ad hoc positioning system (aps). In *Proc. IEEE Globecom*, pages 2926–2931, San Antonio, TX, USA, Nov. 2001.

[3] Y. Shang, W. Ruml, Y. Zhang, and M.P.J. Fromherz. Localization from mere connectivity. In *Proc. ACM MobiHoc*, pages 201–212, Annapolis, Maryland, USA, Jun. 2003.

[4] C. Buschmann, D. Pfisterer, and S. Fischer. Estimating distances using neighborhood intersection. In *Proc. IEEE Emerging Technologies and Factory Automation*, pages 314–321, Prague, Czech, Sep. 2006.

[5] F.L. Villafuerte, K. Terfloth, and J. Schiller. Using network density as a new parameter to estimate distnace. In *Proc. the Seventh International Conference on Networking*, pages 30–35, Cancun, Mexico, Apr. 2008.

[6] D.D. Wackerly, W. Mendenhall, and R.L. Scheaffer. *Mathematical Statistics with Applications*. Suxbury, 2002.

[7] M. Franceschetti and R. Meester. *Random Networks for Communication: From Statistical Physics to Information Systems*. Cambridge University Press, 2007.

[8] S. Mukherjee and D. Avidor. Connectivity and transmit-energy considerations between any pair of nodes in a wireless ad hoc network subject to fading. *IEEE Transactions on Vehicular Technology*, 57(2):1226 –1242, Mar. 2008.

[9] C. Bettstetter and C. Hartmann. Connectivity of wireless multihop networks in a shadow fading environment. *Wirel. Netw.*, 11(5):571–579, 2005.

[10] D. Miorandi and E. Altman. Coverage and connectivity of ad hoc networks in presence of channel randomness. In *Proc. IEEE INFOCOM*, pages 491–502, Miami, FL, USA, Mar. 2005.

[11] T. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall PTR, 2001.

[12] A.J. Coulson, A.G. Williamson, and R.G. Vaughan. A statistical basis for lognormal shadowing effects in multipath fading channels. *IEEE Transactions on Communications*, 46(4):494–502, Apr. 1998.

WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. (2012)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/wcm.2204

RESEARCH ARTICLE

Estimating distances via connectivity in wireless sensor networks

Baoqi Huang[1], Changbin Yu[2,3]*, Brian D. O. Anderson[2] and Guoqiang Mao[4]

[1] Inner Mongolia University, Hohhot, China
[2] Australian National University and NICTA Ltd, Canberra, ACT, Australia
[3] Shandong Computer Science Center, Jinan, China
[4] University of Sydney and NICTA Ltd, Sydney, NSW, Australia

ABSTRACT

Distance estimation is vital for localization and many other applications in wireless sensor networks. In this paper, we develop a method that employs a maximum-likelihood estimator to estimate distances between a pair of neighboring nodes in a static wireless sensor network using their local connectivity information, namely the numbers of their common and non-common one-hop neighbors. We present the distance estimation method under a generic channel model, including the unit disk (communication) model and the more realistic log-normal (shadowing) model as special cases. Under the log-normal model, we investigate the impact of the log-normal model uncertainty; we numerically evaluate the bias and standard deviation associated with our method, which show that for long distances our method outperforms the method based on received signal strength; and we provide a Cramér–Rao lower bound analysis for the problem of estimating distances via connectivity and derive helpful guidelines for implementing our method. Finally, on implementing the proposed method on the basis of measurement data from a realistic environment and applying it in connectivity-based sensor localization, the advantages of the proposed method are confirmed. Copyright © 2012 John Wiley & Sons, Ltd.

KEYWORDS

distance estimation; maximum-likelihood estimator; Cramér-rao lower bound; generic channel model; received signal strength

***Correspondence**

Changbin Yu, RSISE Building 115, Australian National University, Canberra ACT 0200, Australia.
E-mail: Brad.Yu@anu.edu.au

1. INTRODUCTION

Wireless sensor networks, composed of hundreds or thousands of small and inexpensive nodes with constrained computing power, limited memory, and short battery lifetime, can be used to monitor and collect data in a region of interest. Accurate and low-cost node localization is important for various applications in wireless sensor networks, and thus, great efforts have been devoted to developing localization algorithms, including distance-based algorithms and connectivity-based algorithms [1]. Distance-based localization algorithms rely on distance estimates and can achieve relatively good localization accuracy; if distance estimates are unavailable or suffer from huge errors, connectivity-based localization algorithms are applied but generally achieve coarse-grained localization accuracy. Besides, distance estimation is vital for sensor network management, such as topology control [2,3] and boundary detection [4].

In reality, distance estimation can be realized by using information such as received signal strength (RSS), time of arrival (TOA) and time difference of arrival (TDOA) [1]. The RSS method (using RSS measurements) depends on low-cost hardware and only provides coarse-grained distance estimates; by contrast, the TOA and TDOA methods can provide distance estimates with higher accuracy at the cost of extra hardware, but because of cost constraints, it is impractical to equip all sensors in a large-scale sensor network with extra hardware to obtain accurate distance estimates and thus accurate location estimates. Further, although a number of connectivity-based localization algorithms have been proposed, for example, see [5–8], achieving high localization accuracy usually demands a comparatively large number of anchor nodes, hereafter termed simply *anchors*, whose positions are known *a priori* (accordingly, we term other nodes whose positions are not known and need to be determined as *sensors*). In this paper, we shall propose an

attractive distance estimation method that does not rely on extra hardware but provides comparatively accurate distance estimates.

In a static wireless sensor network, two nodes are termed one-hop or immediate neighbors as long as they can directly communicate with each other. An intuitive observation shows that with a higher probability, two geographically close nodes share more common immediate neighbors than two distant nodes. We quantify and exploit this observation to develop a maximum-likelihood estimator (MLE) for estimating the distance between any pair of neighboring nodes on the basis of their local connectivity information. Herein, local connectivity information refers to the numbers of common and non-common immediate neighbors associated with a pair of neighboring nodes. Because only elementary computations and local connectivity information are involved on each node, the proposed method is energy efficient and totally distributed.

In this paper, we present the distance estimation method under a generic channel model, including the ideal unit disk (communication) model and the more realistic log-normal (shadowing) model as special cases (see Section 2 for further details). Then, we take the log-normal model for example to demonstrate the proposed method: the impact of uncertainties in the log-normal model is examined; the bias and standard deviation of derived distance estimates are numerically evaluated; the proposed method, although not comparable with such fine-grained distance estimation techniques as TOA and TDOA, outperforms the well-known RSS method for long distances; the influences of various factors on the problem of estimating distances via connectivity are analyzed on the basis of the Cramér–Rao lower bound (CRLB), and useful guidelines for implementing the proposed method in reality are also derived; and finally, on implementing the proposed method on the basis of measurement data in a realistic environment and also applying it in connectivity-based sensor localization, the advantages of the proposed method are confirmed.

Prior to our work, [9,10] came up with the methods of estimating distances on the basis of the same idea as ours. The neighborhood intersection distance estimation scheme (NIDES) presented in [9] heuristically relates the distance, for example, from node A to node B, to an easily observed ratio, that is, the number of their common immediate neighbors to the number of immediate neighbors of A, and then performs the distance estimation at node A using this ratio and other *a priori* known information. The NIDES assumes the unit disk model and uniformly and randomly deployed wireless sensor networks. Its enhanced version presented in [10] adapted the ratio by taking into account the number of immediate neighbors of node B and heuristically stated that the NIDES could be applied in arbitrary radio models. Although it turns out that the enhanced NIDES leads to the same solution as ours, their entire treatment rests on empirical observations and heuristic formulations rather than theoretical foundations. In comparison with their work, the *contributions* of this paper are as follows: (i) a statistical model is formally established for the distance estimation problem, and an MLE solution with mathematical proofs of the correctness is provided; (ii) the problem is considered under a generic channel model widely used in the literature, including the more realistic log-normal model; (iii) the performance of the proposed method is comprehensively analyzed under the log-normal model in terms of model uncertainties, bias, standard deviation and root mean square error (RMSE); (iv) a CRLB analysis is carried out for the problem of estimating distances via connectivity; and (v) it is shown that the proposed method contributes to the quality of connectivity-based sensor localization.

The remainder of the paper is organized as follows. Section 2 introduces the network model and the (radio) channel model. Section 3 proposes the method under a generic channel model. Under the log-normal model, Section 4 analyzes the performance of the proposed method; Section 5 provides a CRLB analysis for the general distance estimation problem using connectivity; Section 6 implements the proposed method using the measurement data from a real environment; Section 7 reports the contributions of the proposed method to connectivity-based sensor localization by simulations. Finally, we conclude the paper in Section 8.

2. SYSTEM MODEL DESCRIPTION

This section briefly introduces the system model we shall use, including the network model and the channel model. Throughout this paper, we shall use the following mathematical notations: $\Pr\{\cdot\}$ denotes the probability of an event, and $E(\cdot)$ denotes the statistical expectation.

2.1. Network model

In static wireless sensor networks, nodes are often assumed to be randomly and uniformly distributed on account of the random nature of the network deployment. A homogeneous Poisson process provides an accurate model for a uniform distribution of nodes as the network size approaches infinity. Therefore, we consider *a static wireless sensor network that is deployed over an infinite plane according to a homogeneous Poisson process of intensity* λ.

2.2. Channel model

Let P_{T} be the transmitted signal power by a transmitter and $P_{\mathrm{R}}(d)$ be the received signal power by a receiver located at distance d from the transmitter. According to [11], the *log-normal model* predicts $P_{\mathrm{R}}(d)$ to be log-normally distributed and is typically modeled as follows:

$$P_{\mathrm{R}}(d)(\mathrm{dBm}) = \overline{P_{\mathrm{R}}}(d_0)(\mathrm{dBm}) - 10\alpha \log_{10} \frac{d}{d_0} + Z \quad (1)$$

$$\overline{P_{\mathrm{R}}}(d) = \frac{\tau^2 G_{\mathrm{T}} G_{\mathrm{R}} P_{\mathrm{T}}}{(4\pi)^2 d^{\alpha}} \tag{2}$$

where $\overline{P_{\mathrm{R}}}(d_0)$ (dBm) is the mean received signal power in dB m at a reference distance d_0, α is the path loss exponent, G_{T} and G_{R} are the transmitter and receiver antenna gains, τ is the wavelength of the propagating signal in meters, and Z is a random variable representing the shadowing effect, normally distributed with mean zero and variance σ^2_{dB}. Typically, α can vary between 2 in free space and 6 in heavily built urban areas, and σ_{dB} is as low as 4 and as high as 12 according to [11].

If $P_{\mathrm{R}}(d)$ is above some specified value P_{c}, the receiver is able to communicate with the transmitter. Particularly, if $\sigma_{\mathrm{dB}} = 0$, the log-normal model is equivalent to the unit disk model with the transmission range

$$r = \left(\frac{\tau^2 G_{\mathrm{T}} G_{\mathrm{R}} P_{\mathrm{T}}}{(4\pi)^2 P_{\mathrm{c}}} \right)^{\frac{1}{\alpha}} \tag{3}$$

Hence, under the unit disk model, the communication coverage of each node is a perfect disk of radius r. In further discussions, r is not limited to be the transmission range of the unit disk model but a generalized parameter defined by (3).

In effect, the randomness on the received signal power $P_{\mathrm{R}}(d)$ can be described by a function $g(d)$ denoting the probability that a directional communication link exists from transmitter to receiver with distance d. On the basis of $g(d)$, a generic channel model can be defined once $g(d)$ satisfies the following restrictions:

$$\begin{cases} g(d_1) = g(d_2), & \text{if } d_1 = d_2 \quad (4) \\ g(d_1) \le g(d_2), & \text{if } d_1 \ge d_2 \quad (5) \\ 0 < \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\sqrt{x^2 + y^2}) \mathrm{d}x \mathrm{d}y < \infty \quad (6) \end{cases}$$

The generic channel model has been treated intensively in percolation theory [12,13]. The first restriction indicates that the propagation path is symmetric; the second one indicates that $g(d)$ must be a non-increasing function of d; and the third one avoids the trivial cases that any two nodes are directly connected with probability 1 and that any two nodes are isolated with probability 1. It can be easily shown that both the unit disk model and the log-normal model satisfy these restrictions [14].

As transmit power of each node is actually tunable in many wireless sensor networks [15], we require that during the period of running the proposed distance estimation method, all nodes transmit at a *common* power level, that is, P_{T}. Furthermore, throughout the paper, we make the following assumptions (as is commonly the case in the literature).

Assumption 1. *The attenuations caused by shadowing effects (i.e., Z) between any pairs of nodes are independent and identically distributed.*

Assumption 2. *Communication links are symmetric, namely that node A can directly receive packets from node B as long as node B can directly receive packets from node A.*

Even though field measurements in real applications seem to indicate that the attenuations between two links with a common node are correlated [16], Assumption 1 is generally considered appropriate for far field transmission and is widely used in the literature [14,16–20]. Although the above assumptions may not fully reflect a real network environment, they still enable us to obtain some results as estimates for more realistic situations.

3. THE DISTANCE ESTIMATION METHOD

In this section, we present the method of estimating distances via connectivity and detail its implementation under the log-normal model.

3.1. Estimating distances under the unit disk model

In a static wireless sensor network, given two nodes A and B with coordinates $(x_{\mathrm{A}}, y_{\mathrm{A}})$ and $(x_{\mathrm{B}}, y_{\mathrm{B}})$, their distance is defined to be d ($d \le r$), and two disks with the same radius r represent their individual communication coverage under the unit disk model, as shown in Figure 1. Because of $d \le r$, the two disks intersect and create three disjoint regions. Regarding r as a constant, we define $S = \pi r^2$ and $f(d)$ to be the area of the middle region in Figure 1, where

$$f(d) = \frac{2S}{\pi} \arccos\left(\frac{d}{2r}\right) - d\sqrt{r^2 - \frac{d^2}{4}} \tag{7}$$

It is obvious that the nodes residing in the middle region are common immediate neighbors of A and B, the nodes residing in the left (right) one are non-common immediate neighbors of A (B). Define three random variables M, P, and Q to be the numbers of the three categories of neighbors. Obviously, they are mutually independent

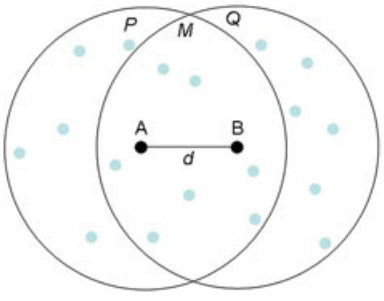


Figure 1. The communication coverage of two nodes under the unit disk model.

and Poisson with means $\lambda f(d), \lambda(S - f(d))$, and $\lambda(S - f(d))$, as pointed out in [13]. The actual values of M, P, and Q can be easily obtained after A and B exchange their neighborhood information. On the basis of the observations of M, P, and Q, an MLE for estimating d is summarized as follows:

Theorem 1. *M, P, and Q are mutually independent Poisson random variables with means $\lambda f(d), \lambda(S - f(d))$, and $\lambda(S - f(d))$, respectively. If $f(d)$ is invertible and S is a non-zero constant, then the MLE for d, termed $\hat{d}$, is*

$$\hat{d} = \begin{cases} f^{-1}(S), & \text{if } M = P = Q = 0; \quad (8) \\ f^{-1}(\hat{\rho}S), & \text{otherwise} \quad (9) \end{cases}$$

where $\hat{\rho} = 2M/(2M + P + Q)$.

Proof. See Appendix A. □

Note that the actual value of λ is not needed in obtaining $\hat{d}$. Although nodes are assumed to follow a random and uniform distribution of density λ (as a result of the Poisson point process), the derivation of the MLE indicates that as long as nodes in a *local* region that covers the communication coverage of two neighboring nodes admits a uniform density, the proposed method is reasonably applicable. Moreover, in some applications, nodes may be placed on the basis of certain regular structures but with noises. For example, in a two-dimensional sensor network, the x-coordinate and y-coordinate of each node are Gaussian with same variance and mean values at one grid point. Compared with a random and uniform distribution, such distribution is even closer to be uniform, and the proposed method can thus attain better performance. In addition, if the least squares method instead of the MLE is applied here, the resulting expression of the distance estimator is actually the same as that in Theorem 1.

3.2. Extension under the generic channel model

Under the generic channel model defined by $g(d)$, M, P, and Q continue to denote the numbers of common and non-common immediate neighbors associated with two nodes. First, we can compute their expectations as follows:

$$E(M + P) = E(M + Q) = \lambda \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \times g(\sqrt{(x - x_B)^2 + (y - y_B)^2}) \mathrm{d}x\mathrm{d}y \quad (10)$$

$$E(M) = \lambda \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\sqrt{(x - x_A)^2 + (y - y_A)^2}) \times g(\sqrt{(x - x_B)^2 + (y - y_B)^2}) \mathrm{d}x\mathrm{d}y \quad (11)$$

Then, from the third restriction on $g(d)$, that is, (6), it follows that $E(M) < E(M + P) < \infty$. Unlike the unit disk model where the independence among M, P, and Q is straightforward because of having three disjoint regions, the generic channel model does not necessarily lead to such three disjoint regions, while the following theorem guarantees the mutual independence.

Theorem 2. *Suppose a static wireless sensor network is deployed in an infinite plane according to a homogeneous Poisson process of density λ and conforms to the generic channel model defined by $g(d)$; given two nodes in this wireless sensor network, let M be the number of their common immediate neighbors and P and Q be the numbers of their non-common immediate neighbors. Then, M, P, and Q are mutually independent Poisson random variables.*

Proof. See Appendix B. □

Under the generic channel model, S and $f(d)$ are generalized to specify the expectations of M, P and Q, instead of the areas defined under the unit disk model, and have the forms of

$$S = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\sqrt{(x - x_B)^2 + (y - y_B)^2}) \mathrm{d}x\mathrm{d}y \quad (12)$$

$$f(d) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\sqrt{(x - x_A)^2 + (y - y_A)^2}) \times g(\sqrt{(x - x_B)^2 + (y - y_B)^2}) \mathrm{d}x\mathrm{d}y \quad (13)$$

Therefore, if S and $f(d)$ satisfy the conditions in Theorem 1, the MLE is directly applicable.

In reality, however, sensor networks are deployed in regions of finite areas, and thus, the expectations of M, P, and Q associated with two nodes, especially those near network boundaries, cannot be derived by simply integrating over an infinite plane to compute S and $f(d)$. This is termed *boundary effects*. In this study, we concentrate on the theoretical foundations of the proposed method and will tackle boundary effects in our future work.

Prior to implementing the proposed method in a static wireless sensor network, it is a premise to know the wireless channel, that is, $g(d)$, such that the quantity S, the function $f(d)$, and its inverse can be determined and then programmed into each node. After that, because of the simple mechanism of the proposed method, a distributed protocol can be easily designed for collecting and exchanging local connectivity information by each node through broadcasting operations. Once a node obtains neighboring information of all its immediate neighbors, this node is able to estimate the distances from its immediate neighbors using the inverse of $f(d)$, S, and the MLE in Theorem 1.

If Assumption 2 holds in a sensor network, each pair of neighboring nodes will have identical information for estimating their distance and thus will obtain the same distance estimate; otherwise, provided that A can hear B but

B cannot hear A, A will estimate their distance, whereas B will not. Such asymmetry in distance estimation can be alleviated by allowing each node to exchange two-hop neighborhood information with its immediate neighbors.

Clearly, if the size of each sensor's neighborhood has the magnitude of $O(1)$, the complexities for communications and computations of running the proposed method in this network are both $O(n)$, where n is the number of nodes in a wireless sensor network, implying that the proposed method is efficient and scalable.

3.3. Implementation under the log-normal model

Provided that the wireless channel is known to be log-normal with known parameters r, σ_{dB}, and α, we shall demonstrate how to determine $g(d)$, S, and $f(d)$ involved in the proposed method.

3.3.1. Formulating $g(d)$

Under the log-normal model, for a transmitter and receiver pair with distance d, if the received signal power described by (1) is no less than P_{c}, a bi-directional communication link exists between them (as a result of the symmetry in Assumption 2). The probability that the two nodes are able to communicate with each other, that is, $g(d)$, is

$$g(d) = \int_{k \log \frac{d}{r}}^{\infty} \frac{\mathrm{e}^{-\frac{z^2}{2\sigma_{\mathrm{dB}}^2}}}{\sqrt{2\pi}\sigma_{\mathrm{dB}}} \mathrm{d}z \tag{14}$$

where $k = 10\alpha/\log 10$ and log denotes natural logarithm. Evidently, $g(d)$ is determined by r, α, and σ_{dB}, where r can be easily computed given the parameters in (3), and α and σ_{dB} can be derived by measurements obtained prior to the deployment of sensor networks or empirically assigned on the basis of the characteristics of the deployment environment [11]. Alternatively, using the technique presented in [21], α and σ_{dB} can be estimated through processing of RSS measurements (and no distance measurements), and depending on the level of noise and amount of measurement data, they may result in imprecise α and σ_{dB}.

We plot $g(d)$ with respect to d and σ_{dB} given $\alpha = 4$ and $r = 1$ in Figure 2(a). It can be seen that the smaller is d, the higher is the probability that a communication link exists; a larger σ_{dB} tends to inhibit communications for a smaller d but promotes communications for a larger d in comparison with a smaller σ_{dB}.

In view of the restrictions on $g(d)$, it is straightforward to obtain $\lim_{d\to\infty} g(d) = 0$; as such, for an extremely small and positive ε, there exists d_{th} such that $g(d) < \varepsilon$ if $d > d_{\mathrm{th}}$. That is to say, nodes with distances to a node longer than d_{th} hardly communicate with this node directly; as such, d_{th} is a surrogate of the transmission range. This phenomenon can be observed in Figure 2(a).

3.3.2. Formulating S and $f(d)$

In [17,19], the expectations $E(M+P)$ (or $E(M+Q)$) has been well studied. Thus, we can have

$$S = \pi r^2 \mathrm{e}^{\frac{2\sigma_{\mathrm{dB}}^2}{k^2}} \tag{15}$$

By (11), (13), and (14), we can derive the formula for $f(d)$ under the log-normal model. By letting $\alpha = 4$ and $r = 1$, we plot $f(d)$ with respect to different values of d and σ_{dB} in Figure 2(b).

As can be seen in Figure 2(b), $f(d)$ is monotonically decreasing and invertible; hence, Theorem 1 is applicable under the log-normal model. But the closed-form formula for $f(d)$ and its inverse are unavailable. Alternatively, we can establish a piecewise linear function to approximate its inverse; for each affine segment, a linear regression model is applied to predict d. Considering the fact that two nodes with distance longer than d_{th} hardly communicate with each other directly, we restrict the distance estimates to be between 0 and d_{th}. But in a real estimation process, $\hat{\rho}S$ and S may exceed $[f(d_{\mathrm{th}}), f(0)]$ and consequently $\hat{d}$ may exceed $[0, d_{\mathrm{th}}]$. Therefore, we can obtain the distance estimator as follows:

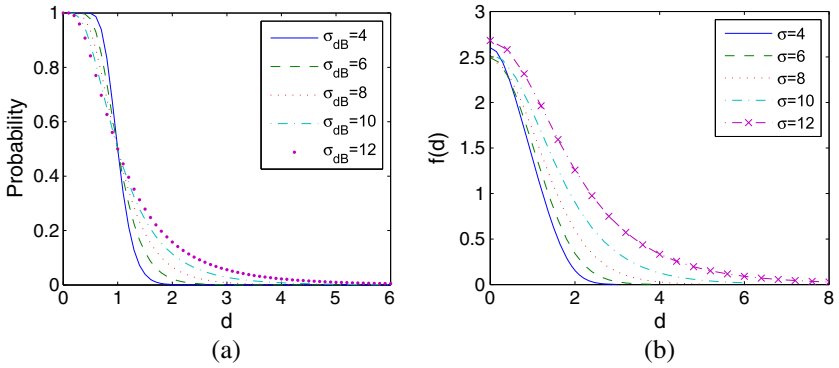


Figure 2. The functions $g(d)$ and $f(d)$ under the log-normal model with $\alpha = 4$ and $r = 1$.

$$\hat{d} = \begin{cases} 0, & \text{if } M = P = Q = 0 \text{ or } \hat{\rho}S > f(0); \\ f^{-1}(\hat{\rho}S), & \text{if } f(d_{\text{th}}) \le \hat{\rho}S \le f(0); \\ d_{\text{th}}, & \text{if } \hat{\rho}S < f(d_{\text{th}}) \end{cases} \tag{16}$$

4. PERFORMANCE ANALYSIS

In this section, we evaluate the performance of the proposed method under the log-normal model from different respects.

4.1. Impact of imprecise α and σ_{dB}

In the proposed method, the parameters of the log-normal model, that is, α and σ_{dB}, are supposed to be known precisely. But what if their values are imprecise? To answer this question, we define

$$\rho_{\alpha,\sigma_{\text{dB}}}(d) = \frac{f(d)}{S} \tag{17}$$

where $f(d)$ and S are computed using (13), (15), and (14) given α and σ_{dB}. Thus, the distance estimator in (16) is equivalently

$$\hat{d} = \begin{cases} 0, & \text{if } M = P = Q = 0 \text{ or } \hat{\rho} > \rho_{\alpha,\sigma_{\text{dB}}}(0); \\ \rho^{-1}_{\alpha,\sigma_{\text{dB}}}(\hat{\rho}), & \text{if } \rho_{\alpha,\sigma_{\text{dB}}}(d_{\text{th}}) \le \hat{\rho} \le \rho_{\alpha,\sigma_{\text{dB}}}(0); \\ d_{\text{th}}, & \text{if } \hat{\rho} < \rho_{\alpha,\sigma_{\text{dB}}}(d_{\text{th}}) \end{cases} \tag{18}$$

Evidently, using imprecise α and/or σ_{dB} results in an incorrect function $\rho^{-1}_{\alpha,\sigma_{\text{dB}}}(\hat{\rho})$, with the result that the accuracy of the distance estimate $\hat{d}$ is degraded. We plot the function $\rho^{-1}_{\alpha,\sigma_{\text{dB}}}(\hat{\rho})$ with respect to different values of α and σ_{dB} in Figure 3.

Supposing σ_{dB} is known to be exactly 4, we investigate the impact of the uncertainty in α. As shown in Figure 3(a), $\rho^{-1}_{\alpha,\sigma_{\text{dB}}}(\hat{\rho})$ is much more sensitive to a small α than to a large α; in other words, for a small α, using an imprecise version of α tends to degrade the accuracy of the distance estimate $\hat{d}$ more seriously than for a large α. Moreover, if α is overestimated, then an underestimated $\hat{d}$ will be produced for a small d but an overestimated $\hat{d}$ for a large d, and vice versa. However, $\rho^{-1}_{\alpha,\sigma_{\text{dB}}}(\hat{\rho})$ does not demonstrate the same sensitivity to σ_{dB} as is observed for α, as illustrated in Figure 3(b). We can conclude that if σ_{dB} is overestimated, then an overestimated $\hat{d}$ will be produced for a small d but an underestimated $\hat{d}$ for a large d, and vice versa.

4.2. Bias and standard deviation

According to Theorem 1, all possible values of $\hat{\rho}$ are rational numbers within $[0, 1]$ so that $\hat{d}$ is a discrete random variable and its jth moment is as follows:

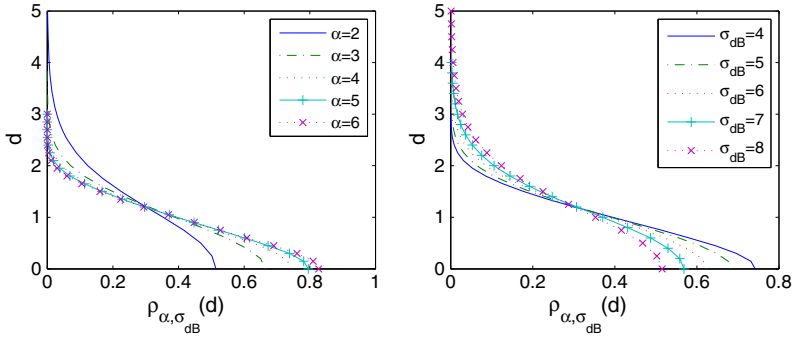
$$E(\hat{d}^j) = \sum_a \left[\hat{d}^j \Pr(\hat{d} = a)\right] \tag{19}$$

We divide the range of $\hat{d}$, that is, $[0, d_{\text{th}}]$, into w equal intervals: $\mathcal{I}_1 = [z_0, z_1), \ldots, \mathcal{I}_w = [z_{w-1}, z_w]$ with $z_i = (i \times d_{\text{th}})/w$. Given a sufficient large w, $\hat{d}$ is approximately constant over each interval, denoted $\tilde{d}_i$. Then we can approximately reformulate (19) as

$$E(\hat{d}^j) \approx \sum_{i=1}^{w} \left[(\tilde{d}_i)^j \Pr(\hat{d} \in \mathcal{I}_i)\right] \tag{20}$$

Toward the probability associated with the ith interval $\mathcal{I}_i$, we have

$$\Pr(\hat{d} \in \mathcal{I}_i) = \begin{cases} \Pr(f(z_1) < \hat{\rho}S \le S)), & \text{if } i = 1; \\ \Pr(f(z_i) < \hat{\rho}S \le f(z_{i-1})), & \text{if } 1 < i < w; \\ \Pr(0 \le \hat{\rho}S \le f(z_{w-1})), & \text{if } i = w \end{cases}$$



(a) $\sigma_{dB} = 4$ (b) $\alpha = 4$

Figure 3. The inverse function of $\rho_{\alpha,\sigma_{\text{dB}}}(d)$ under the log-normal model.

Table I. The values of λ with respect to different σ_{dB} and μ when $\alpha = 4$.

	μ				
σ_{dB}	5	10	20	30	40
0	1.59	3.18	6.37	9.55	12.73
4	1.43	2.86	5.73	8.59	11.45
8	1.04	2.08	4.17	6.25	8.33
12	0.61	1.23	2.45	3.68	4.90

By letting $Y = P + Q$, we have

$$\Pr(b < \hat{\rho}S < c) = \sum_{y=0}^{\infty} [\Pr(b < \hat{\rho}S < c | Y = y) \times \Pr(Y = y)] \tag{21}$$

which makes it possible for us to numerically evaluate the moments of $\hat{d}$ and thus the bias and standard deviation.

Let μ be the expected number of immediate neighbors of a node, namely $\mu = E(M + P) = E(M + Q)$, and the values of λ with respect to different σ_{dB} and μ are listed in Table I. For better presentation, the connectivity index μ will be used in the following discussions instead of the node density λ.

Given $\alpha = 4$, $r = 1$, $w = 1000$, and μ varying from 5 to 40, Figure 4 depicts the numerical bias and standard deviation associated with the proposed method and the corresponding simulation results. The two groups of results are highly consistent, and the comparatively non-smooth aspect of some of the curves, for example, Figure 4(a), is probably attributable to the fact that all observations of M, P, and Q in related results are necessarily integer, and such observations are used in determining the curves.

As shown in Figure 4, the proposed method is obviously *biased*, but the absolute bias is much less than the standard deviation in most cases; except for $\sigma_{\text{dB}} = 0$, the absolute bias and standard deviation are comparable with true distances, especially for short distances and sparse sensor networks, and particularly, when $\mu = 5$, their values are extraordinarily large and nearly twice the corresponding values when $\mu = 10$. Moreover, with μ increasing, the standard deviation always reduces, while the absolute bias reduces in most cases. An intuitive explanation is that with μ increasing, the variances of the ratios $2M/E(2M)$ and $(2M+P+Q)/E(2M+P+Q)$ both decrease, the variance of $\hat{\rho}$ is reduced and so is the variance of $\hat{d}$. As mentioned in the previous section, a large σ_{dB} promotes communications between distant nodes but inhibits communications between close nodes; as a result, connectivity is related to a wide range of distances so that the geometric information implied by connectivity becomes less accurate. Hence, the larger is the σ_{dB}, the worse are both the bias and the standard deviation.

4.3. Root mean square error

As a performance measure, the RMSE is defined to be the square root of the sum of the square bias and variance of estimation errors. We plot the RMSE of $\hat{d}$ produced by the proposed method in Figure 5. As can be seen, the RMSE decreases with μ increasing and σ_{dB} decreasing, which is consistent with how the bias and standard deviation of $\hat{d}$ depend on μ and σ_{dB}. When d is near 0, the RMSE is extraordinarily large compared with the true value of d, implying that the proposed method fails to provide reasonable estimates for short distances. This underperformance with short distances limits the use of the proposed method in practice and is due to a mixed impact of the following facts:

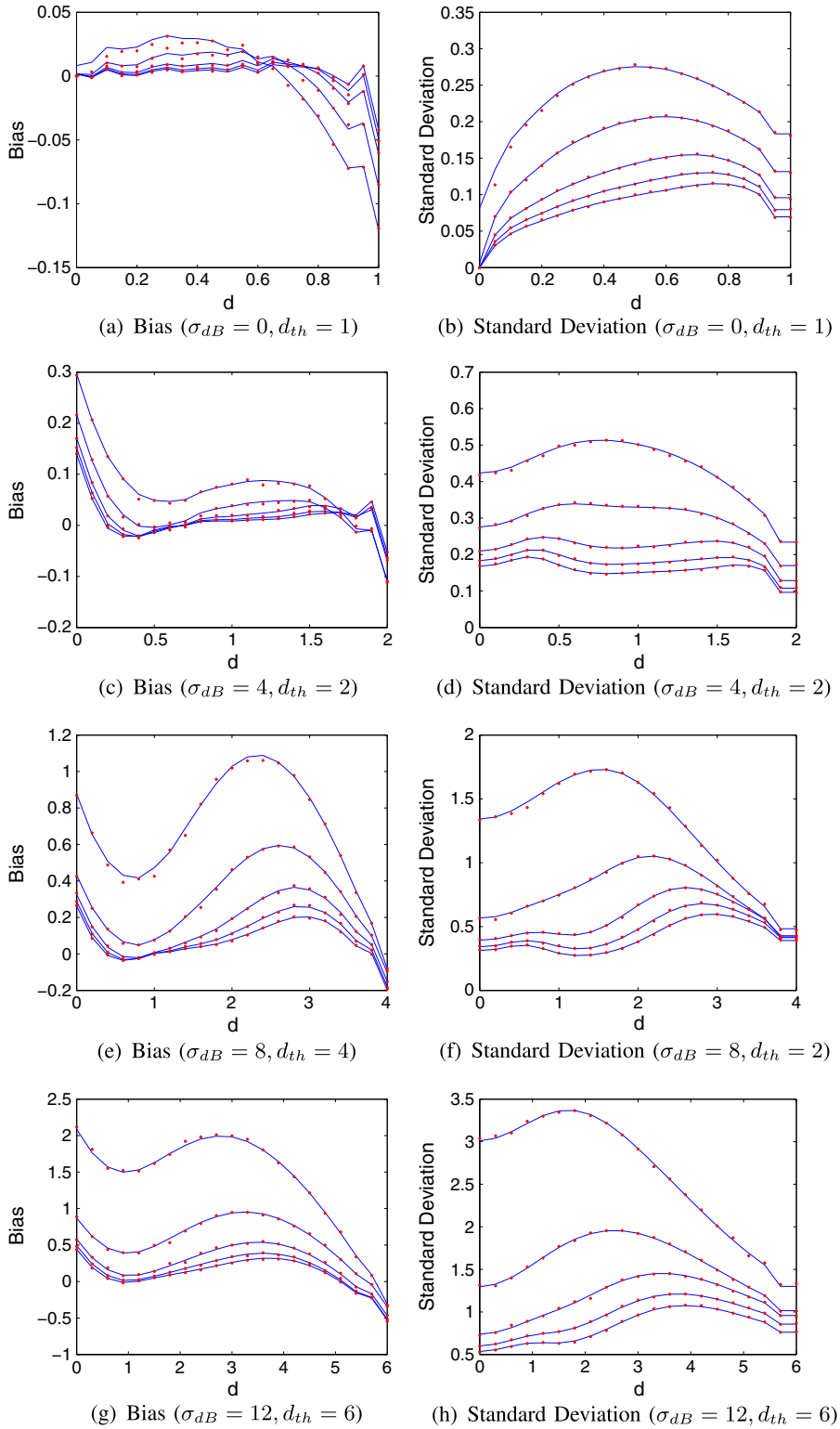
- It is evident that the variance of $2M + P + Q$ is constant no matter what d is, but the variance of $2M$ increases with d decreasing; as a result, $\hat{\rho}$, that is, $2M/(2M + P + Q)$, is more likely to suffer bigger variances when d is small than when d is large.
- As depicted in Figure 3, $\rho^{-1}_{\alpha,\sigma_{\text{dB}}}(\hat{\rho})$ is quite sensitive to $\hat{\rho}$ when d is small, namely that a small perturbation in $\hat{\rho}$ leads to a big change in $\hat{d}$ and thus a big distance estimation error.
- In light of (18), $\hat{d}$ is roughly set 0 when $\hat{\rho}$ is greater than $\rho_{\alpha,\sigma_{\text{dB}}}(0)$, but a small d often causes $\hat{\rho}$ to be within $[\rho_{\alpha,\sigma_{\text{dB}}}(0), 1]$, and hence the underperformance is attained.

To conclude, for short distances, the non-smooth aspect and the sensitivity to $\hat{\rho}$ of the function defined in (18) are responsible for the underperformance.

Under the log-normal model with $\sigma_{\text{dB}} > 0$, distance estimation can be realized by using the RSS measurements, that is, received signal powers. The bias and variance of the resulting distance estimate (denoted $\hat{d}_{\text{RSS}}$) are provided in [22] so that we can compute the RMSE of $\hat{d}_{\text{RSS}}$ and compare it with that of $\hat{d}$ in Figure 5. It can be seen that (i) the RMSE of $\hat{d}_{\text{RSS}}$ increases in direct proportion to d, but that of $\hat{d}$ appears to have comparatively small variations with d increasing, and (ii) the proposed method outperforms the RSS method for long distances by a large margin.

5. ANALYSIS BASED ON THE CRAMÉR–RAO LOWER BOUND

In this section, we formulate the CRLB regarding the distance estimation problem via connectivity, that is, estimating d from M, P, and Q, under the log-normal model. For this estimation problem, the unknown parameters are d and λ. The Fisher Information Matrix (FIM) for this estimation problem, denoted FIM(d, λ), is



(a) Bias ($\sigma_{dB} = 0, d_{th} = 1$) (b) Standard Deviation ($\sigma_{dB} = 0, d_{th} = 1$)

(c) Bias ($\sigma_{dB} = 4, d_{th} = 2$) (d) Standard Deviation ($\sigma_{dB} = 4, d_{th} = 2$)

(e) Bias ($\sigma_{dB} = 8, d_{th} = 4$) (f) Standard Deviation ($\sigma_{dB} = 8, d_{th} = 2$)

(g) Bias ($\sigma_{dB} = 12, d_{th} = 6$) (h) Standard Deviation ($\sigma_{dB} = 12, d_{th} = 6$)

Figure 4. The bias and standard deviation of distance estimation from numerical evaluations (solid lines) and simulations (dashed lines) with $\mu = 5, 10, 20, 30$, and 40, $\alpha = 4$, and $r = 1$. For the standard deviation, a larger μ corresponds to a line to the bottom; for the bias, a larger μ corresponds to a line with the bias closer to 0.

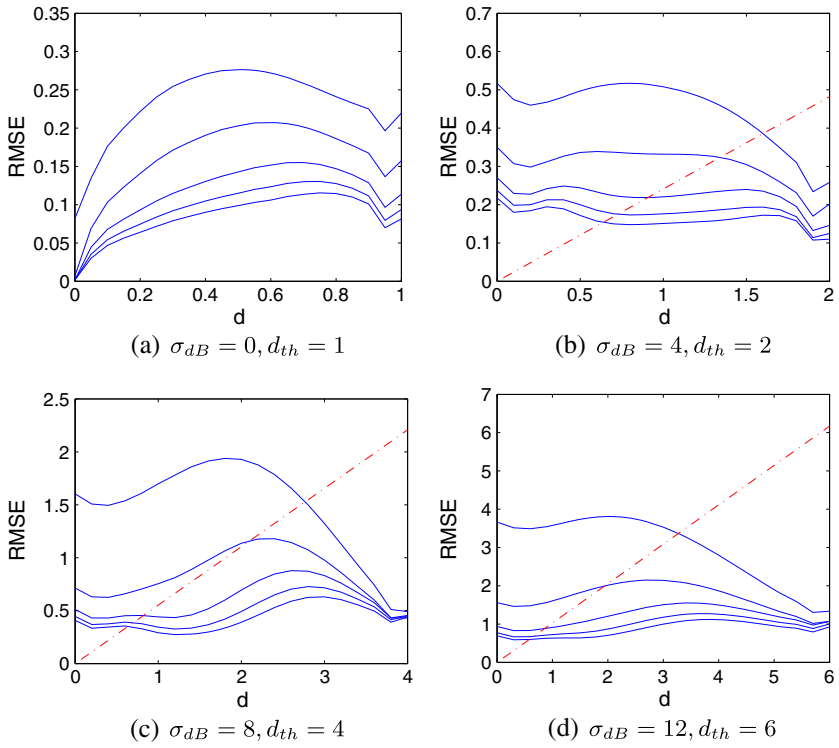


Figure 5. The RMSE of $\hat{d}_{\mathrm{RSS}}$ (dashed lines) and $\hat{d}$ (solid lines; $\mu = 10, 20, 30$, and 40 with $\alpha = 4$ and $r = 1$; a larger μ corresponds to a curve to the bottom) under the log-normal model.

$$\mathrm{FIM}(d,\lambda) = \begin{pmatrix} \lambda\left(f'(d)\right)^2\left(\frac{1}{f(d)} + \frac{2}{S-f(d)}\right) & -f'(d) \\ -f'(d) & \frac{2S-f(d)}{\lambda} \end{pmatrix} \tag{22}$$

where $f(d)$ is differentiable to the first order in (13) (see Appendix C). Then, the CRLB for d by using any unbiased estimator, denoted $\mathrm{CRLB}(d)$, is

$$\mathrm{CRLB}(d) = \frac{(S - f(d))(2S - f(d))f(d)}{2\lambda S^2 (f'(d))^2} \tag{23}$$

Although the CRLB is only valid for unbiased distance estimates and the proposed method is known to be biased, it is still helpful to understand the essential features of the distance estimation problem. In what follows, we shall investigate the influences of various parameters.

5.1. Influence of λ

It is clear that the CRLB is inversely proportional to λ. In other words, better estimation accuracy can be attained in dense wireless sensor networks, which is intuitive and is also illustrated in Figure 5. Hence, it is attractive to apply the proposed method in dense wireless sensor networks. Dense wireless sensor networks, however, are really required in some circumstances. For example, because of the limited energy resource at each node, nodes are usually deployed in high density and may take turns to be active in order to prolong the network lifetime [23]; accordingly, many scheduling strategies have been developed to determine when and which sensors should be powered up and which sensors should be put into energy saving mode while satisfying certain coverage and connectivity requirements [24–29].

5.2. Influence of d

According to (23), it is difficult to directly observe the influence of d on the CRLB, for we do not have the closed-form formulas for $f(d)$ and $f'(d)$ except for the case of $\sigma_{\mathrm{dB}} = 0$. But because it is easily justified that the numerator of (23) is bounded in a narrow range, if the denominator can be very small, the CRLB will be seriously affected by the denominator. On the basis of Figure 2(b), we can obtain some preliminary knowledge about the key component in the denominator, that is, $f'(d)$. As can be seen from Figure 2(b), with d increasing from 0, $|f'(d)|$ firstly experiences a rise and then decreases after d is greater than some value that differs from σ_{dB}; when d increases further, $|f'(d)|$ continuously decreases and approaches 0. Hence, it is postulated that the CRLB will experience a rise with d increasing.

As shown in Figure 4, in the cases of $\sigma_{\mathrm{dB}} = 8$ and 12, the standard deviation displays an evident rise when d is

larger than some value and then drops when d approaches d_{th}. The reason causing such a drop is that we restrict the maximal distance estimate to be d_{th} so that estimates for d near d_{th} are improved. In the case of $\sigma_{\text{dB}} = 4$, $|f'(d)|$ is not so close to 0 when d is near d_{th} but is comparatively small when d is near 0; as a result, the expecting rise does not happen.

5.3. Influences of P_{T} and P_{c}

Provided that G_{T}, G_{R}, and α are known in (3), r is proportional to $(P_{\text{T}}/P_{\text{c}})^{\frac{1}{\alpha}}$, where P_{T} is the transmission power and P_{c} is the threshold of power for communications. If both P_{T} and P_{c} are tunable in wireless sensor networks (which does not break requirement on the common transmission power P_{T}), it will be meaningful to analyze their influences on the CRLB. For simplicity, we shall use r instead of P_{T} and P_{c} to carry out the analysis. To do so, we derive the following theorem.

Theorem 3. *Consider the CRLB that is defined on the basis of (13), (15), and (23), and suppose only d and r are variables. Then, the CRLB for a given distance d with $r = r_0$ is equal to the CRLB for a distance d/r_0 with $r = 1$.*

Proof. See Appendix D. □

This theorem reveals that (i) the CRLB is virtually determined by the ratio d/r and (ii) the value field of the CRLB is invariant no matter how large r is. That is, if the ratio $P_{\text{T}}/P_{\text{c}}$ is raised, distant nodes will tend to become connected so that estimates for long distances will be available, but the CRLB will not exceed the value field of the CRLB associated with the original small value of $P_{\text{T}}/P_{\text{c}}$. Consequently, estimates for long distances will generally suffer less relative errors (i.e., the ratio of the estimation error to d) than those for short distances.

Notice that because the CRLB is not monotonic with d, tuning $P_{\text{T}}/P_{\text{c}}$ does not definitely increase or decrease the corresponding CRLB associated with one given value of d. Because raising $P_{\text{T}}/P_{\text{c}}$ results in more immediate neighbors for each node and consequently more distance estimates, although any distance estimate is not necessarily improved, more available distance estimates will benefit other applications, for example, sensor network localization.

Furthermore, this feature can be exploited in the implementation of the proposed method. Considering the fact that in static wireless sensor networks, the procedure of estimating distances is usually executed only once, and probably in the beginning of the network lifetime, the ratio $P_{\text{T}}/P_{\text{c}}$ can be initially set a high value to achieve a high 'sensor density' by increasing P_{T} and/or decreasing P_{c} and then is tuned to be a normal value after the phase of estimating distances. As a result, more estimates of long distances with comparatively good accuracies will be available.

6. IMPLEMENTING THE METHOD IN PRACTICE

In this section, we improve the proposed method when dealing with short distances and then test it in a practical environment.

6.1. Dealing with short distances

Given α and σ_{dB}, define $\epsilon_{\alpha,\sigma_{\text{dB}}}$ to be the RMSE when $d = 0$. As illustrated in Figure 5, the RMSE of distance estimates produced by the proposed method experiences small variations as d increases from 0 up to d_{th}, so that if $d \geq \epsilon_{\alpha,\sigma_{\text{dB}}}$ the RMSE tends to be under d, implying that relatively good performance is attained. Moreover, on the grounds of the analysis in Section 4.3, we focus on the function defined by (18) with $\hat{\rho} \in [\rho_{\alpha,\sigma_{\text{dB}}}(\epsilon_{\alpha,\sigma_{\text{dB}}}), 1]$ and reformulate it by a linear function $[(1-\hat{\rho})\epsilon_{\alpha,\sigma_{\text{dB}}}]/[1-\rho_{\alpha,\sigma_{\text{dB}}}(\epsilon_{\alpha,\sigma_{\text{dB}}})]$ that smoothly transforms any $\hat{\rho}$ between $\rho_{\alpha,\sigma_{\text{dB}}\text{dB}}(\epsilon_{\alpha,\sigma_{\text{dB}}})$ and 1 to a distance estimate between 0 and $\epsilon_{\alpha,\sigma_{\text{dB}}}$. Consequently, (18) is updated to be

$$\hat{d} = \begin{cases} 0, & \text{if } M = P = Q = 0; \\ \frac{(1-\hat{\rho})\epsilon_{\alpha,\sigma_{\text{dB}}}}{1-\rho_{\alpha,\sigma_{\text{dB}}}(\epsilon_{\alpha,\sigma_{\text{dB}}})}, & \text{if } \hat{\rho} > \rho_{\alpha,\sigma_{\text{dB}}}(\epsilon_{\alpha,\sigma_{\text{dB}}}); \\ \rho^{-1}_{\alpha,\sigma_{\text{dB}}}(\hat{\rho}), & \text{if } \rho_{\alpha,\sigma_{\text{dB}}}(d_{\text{th}}) \leq \hat{\rho} \\ & \quad \leq \rho_{\alpha,\sigma_{\text{dB}}}(\epsilon_{\alpha,\sigma_{\text{dB}}}); \\ d_{\text{th}}, & \text{if } \hat{\rho} < \rho_{\alpha,\sigma_{\text{dB}}}(d_{\text{th}}) \end{cases} \tag{24}$$

6.2. Test the method in a real environment

In [30], a sensor network consisting of 44 nodes was deployed in a real environment, and RSS measurements between any two nodes were reported. On the basis of their measurement data, we can simulate a realistic environment to implement our method. According to [30], $\alpha = 2.3$, $\sigma_{\text{dB}} = 3.92$, and $\overline{P_{\text{R}}}(R_0) = -37.47$ dBm. But to proceed with the experiment, we also need to specify the threshold power P_{c}, which essentially defines whether two nodes are 'connected', and $\epsilon_{\alpha,\sigma_{\text{dB}}}$ in (24). After that, we can compute the function $g(d)$ associated with this channel and then obtain the distance estimators on the basis of (18) and (24), respectively. To avoid boundary effects as much as possible, we consider the four nodes near the center of the deployment region, that is, nodes 15, 23, 24, and 25, and only estimate the inter-node distances associated with the four nodes by using the originally proposed method and the method with the adjustment.

In this experiment, by letting $\epsilon_{\alpha,\sigma_{\text{dB}}}$ be $0.5r$ and raising P_{c} from -61 to -52 dBm, the average distance estimation errors incurred by the original and adjusted methods are listed in Table II. By the distance estimates produced by the RSS method (which were also provided by [30]), we compute the corresponding average distance estimation error, that is, 1.07 m. As depicted in the table, (i) the adjusted method always outperforms the original method; (ii) the

Table II. The experimental results with respect to different P_c.

P_c (dB m)	r (m)	ND	DEE (m)	DEE adjusted (m)
−52	4.28	8.50	1.00	0.44
−53	4.73	11.25	1.49	0.56
−54	5.23	14.00	1.95	0.95
−55	5.78	17.25	1.78	0.74
−56	6.38	20.75	1.81	0.70
−57	7.05	27.25	2.00	1.08
−58	7.80	31.75	2.08	1.41
−59	8.62	34.50	2.11	1.57
−60	9.52	37.00	2.15	1.72
−61	10.53	38.75	2.14	1.67

ND, average node degree; DEE, average distance estimation error.

original method outperforms the RSS method only in the case of $P_c = -52$ dBm, whereas the adjusted method outperforms the RSS method if P_c (dBm) is between −56 and −52 dBm; and (iii) although the average node degree increases with P_c decreasing, the average error obtained with the adjusted method increases in general, a phenomenon which is attributable to boundary effects.

7. APPLYING THE METHOD IN LOCALIZATION

In this section, we report the improvement in connectivity-based sensor localization by using the proposed method.

7.1. Connectivity-based localization algorithms

Connectivity-based localization algorithms, for example, [6,8], generally involve as a crucial component a mechanism of converting connectivity information into rough distance estimates, which are then used for localization as in distance-based localization algorithms.

The DV-hop scheme proposed in [6] employs distance vector exchange. Both sensor and anchor exchange distance tables that contain the locations of and the hop counts to anchors with their corresponding neighboring nodes. Once an anchor obtains the distance tables from other anchors, it estimates an individual average distance per hop and broadcasts this average distance into the network. A sensor approximates its geographic distance to an anchor by multiplying the hop count to this anchor and the associated average distance per hop and then estimates its location by performing trilateration if a sufficient number of distance estimates are obtained. Its variant, that is, the DV-distance scheme, is almost the same as the DV-hop scheme except that it employs the geographic distances measured with the use of radio signals other than hop counts.

Multi-dimensional scaling map (MDS-MAP) proposed in [8] approximates the distance between two nodes by the length of their shortest path and then uses multidimensional scaling to generate a relative map that represents the relative positions of nodes. Once a sufficient number of anchors are known, MDS-MAP estimates the absolute coordinates of all the sensors in the network. Like DV-distance, MDS-MAP can also employ geographic distance measurements; we term it MDS-MAP distance.

In both DV-hop and MDS-MAP, the distance between two nodes is roughly estimated according to the length of the shortest path between them, namely that the one-hop distances along any shortest path are assumed to be equal. As opposed to this assumption, our method provides comparatively accurate estimates of one-hop distances and thus helps to improve the quality of connectivity-based sensor localization, which will be demonstrated in the following subsection.

Many methods have been developed in the literature to improve DV-hop. For instance, in [31], estimating the distance from a sensor to an anchor not only uses the length of the shortest path between them as in DV-hop but also exploits the lengths of the shortest paths from this sensor's immediate neighbors to the anchor. Moreover, DV-hop suffers large errors in anisotropic networks, because the estimates of distances from sensors to anchors can be extraordinarily inaccurate. Accordingly, [32–34] were developed to alleviate the impacts of the anisotropic network topology on the estimates of distances. Because comparatively accurate estimates of one-hop distances provided by our method are the basis for estimating distances from sensors to anchors, it is attractive to combine our method with these DV-hop related methods to improve the estimates of distances from sensors to anchors and thus to improve localization accuracy.

7.2. Simulations

We simulate connectivity-based sensor localization under the log-normal model using DV-hop and MDS-MAP and their distance-based counterparts DV-distance and MDS-MAP distance (with distance measurements from the adjusted method).

To avoid boundary effects, we actually generate wireless sensor networks over a large square with side of 18 but only localize the nodes inside of a small one with side of 6 and concentric to the large one. However, the nodes outside of the small one are sometimes used in estimating distances between nodes inside. Four nodes inside the small square and closest to its four corners are chosen as anchors. Regarding the constants parameterizing the log-normal model, α is known to be 4, σ_{dB} takes values from 0, 4, 8, 12, and P_T and P_c are properly assigned such that $r = 1$. Furthermore, λ takes proper values such that μ varies from 10 to 40 with step size 5 (see Table I).

For each choice of σ_{dB} and μ, 100 independent runs are carried out. In each run, first, a static wireless sensor network is generated according to a homogeneous Poisson process of density λ; second, distance between any pair of neighboring nodes is estimated on the basis of Theorem 1, and accordingly, the average absolute distance estimation

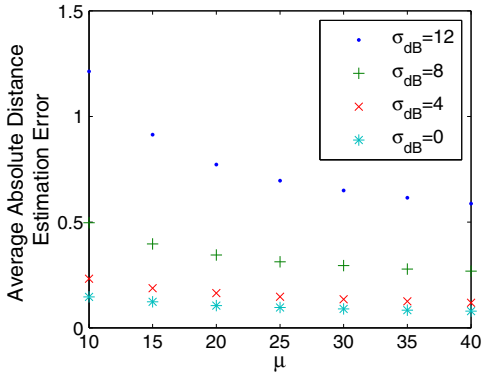


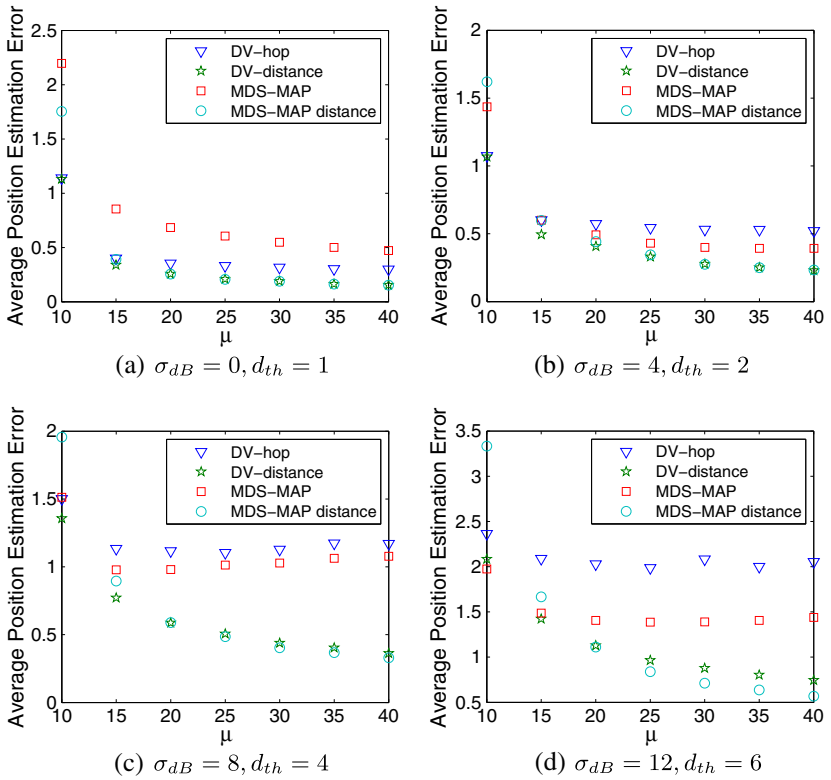
Figure 6. Average absolute distance estimation errors with $\alpha = 4$ and $r = 1$.

error is computed; third, sensors are localized by using four localization algorithms: DV-hop, MDS-MAP, and their distance-based counterparts DV-distance and MDS-MAP distance (with distances coming from the second step); and finally, the average absolute distance estimation error and the average position estimation error are computed for each localization algorithm. Then, the average absolute distance estimation errors and the average position estimation errors are averaged over the 100 independent runs and plotted in Figures 6 and 7.

DV-hop and DV-distance are almost the same except that DV-hop assumes identical one-hop distances along any shortest path, whereas DV-distance uses distance estimates with comparatively good accuracies; this is also true for MDS-MAP and MDS-MAP distance. Because only connectivity information with the assistance of four anchors is exploited to realize sensor localization in the simulations, the fact that DV-distance and MDS-MAP distance use distance estimates produced by our method and their superior performance imply the advantages of our method.

8. CONCLUSIONS

In this paper, we proposed the method of estimating distances via connectivity in static wireless sensor networks by dealing with a generic channel model, including the realistic log-normal model. The proposed method is not relying on extra hardware, totally distributed and energy efficient due to its simple mechanism and computations. Under the log-normal model, the bias and standard deviation of distance estimates from the proposed method were numerically evaluated and verified by simulations; the proposed method outperforms the RSS method for long distances; a CRLB analysis was carried out for the problem of estimating distances using connectivity, and useful guidelines for implementing the proposed method were derived; the influences of uncertainties in the log-normal



(a) $\sigma_{dB} = 0, d_{th} = 1$ (b) $\sigma_{dB} = 4, d_{th} = 2$

(c) $\sigma_{dB} = 8, d_{th} = 4$ (d) $\sigma_{dB} = 12, d_{th} = 6$

Figure 7. Average position estimation errors with $\alpha = 4$ and $r = 1$.

model were examined; and finally, extensive simulations confirmed the advantages of the proposed method. In future work, we may tackle estimating distances involving nodes near the network boundaries, implement the method in a realistic environment and combine the method with the DV-hop related methods to make further improvement.

APPENDIX A: PROOF OF THEOREM 1

Establish a statistical model: observations of M, P and Q provide measured data $\phi = [\,m \quad p \quad q\,]$ where m, p, q are non-negative integers; the unknown parameters are $\theta = [\,d \quad \lambda\,]$. By formulating the likelihood function, we can derive that the MLE for d is the solution for d in the following equation set:

$$\begin{cases} \dfrac{m}{f(d)} - \dfrac{p+q}{S-f(d)} + \lambda = 0, & (25) \\ \dfrac{m+p+q}{\lambda} - (2S - f(d)) = 0 & (26) \end{cases}$$

By eliminating λ, we can obtain

$$2mS = (2m + p + q) f(d) \tag{27}$$

If $2m + p + q > 0$, $\hat{d} = f^{-1}\{[2m/(2m + p + q)]S\}$; otherwise, the solution for d is not well defined, but because $f(d) = S$ maximizes the likelihood, we have $\hat{d} = f^{-1}(S)$. Thus, we prove the theorem.

APPENDIX B: PROOF OF THEOREM 2

Let the two nodes be A and B and consider a finite region of area D covering A and B. As pointed out in [13], apart from A and B, the rest of the Poisson process is not affected, namely that the number of remaining nodes in this region, denoted N, is still Poisson with mean λD. Choosing an arbitrary node C from the N nodes, one of the following cases hold: (i) C can directly communicate with A and B; (ii) C can directly communicate with A but not B; (iii) C can directly communicate with B but not A; and (iv) C cannot directly communicate with A or B.

Conduct a trial for each of the N nodes to decide how it communicates with A and B, and each trial results in the above four cases with probabilities p_1, p_2, p_3, and p_4, respectively. Because of the independence of connectivity assumed in Assumption 1, the N trials are then independent from each other. Evidently, M, P, and Q represent the numbers of nodes belonging to the first three cases. In addition, let the random variable L denote the number of nodes belonging to the last case. Because of $p_1 + p_2 + p_3 + p_4 = 1$, M, P, Q, and L follow a multinomial distribution with parameters N and p_1, p_2, p_3, and p_4. Considering N is Poisson with mean λD, from the theorem on page 8 in [13], it follows that M, P, Q, and L are mutually independent Poisson random variables with means $\lambda D p_1, \lambda D p_2, \lambda D p_3$, and $\lambda D p_4$, respectively. Now, we let the region approach the infinite plane and can conclude that M, P, and Q are mutually independent Poisson with finite means.

APPENDIX C: THE EXISTENCE OF $f'(d)$

At first, consider the following expression:

$$\begin{aligned} &\lim_{\varepsilon\to 0}\left(-\frac{1}{\varepsilon}\left[g(\sqrt{x^2+d^2-2xd\cos\theta}) \right.\right. \\ &\qquad \left.\left. -g(\sqrt{x^2+(d+\varepsilon)^2-2x(d+\varepsilon)\cos\theta})\right]\right) \\ &= \frac{k(x\cos\theta - d)\mathrm{e}^{\frac{-(k(\log\sqrt{x^2+d^2-2xd\cos\theta}-\log r))^2}{2\sigma_{\mathrm{dB}}^2}}}{\sqrt{2\pi}\sigma_{\mathrm{dB}}(x^2+d^2-2xd\cos\theta)} \end{aligned}$$

which is bounded for $x \in [0, +\infty)$. Moreover, the derivative of $f(d)$ can be formulated as

$$\begin{aligned} f'(d) = \int_0^\infty \int_0^{2\pi} &\times g(x)x \lim_{\varepsilon\to 0}\left(-\frac{1}{\varepsilon}\left[g(\sqrt{x^2+d^2-2xd\cos\theta}) \right.\right. \\ &\left.\left. -g(\sqrt{x^2+(d+\varepsilon)^2-2x(d+\varepsilon)\cos\theta})\right]\right)\mathrm{d}\theta\mathrm{d}x \end{aligned} \tag{28}$$

Because $\int_0^\infty \int_0^{2\pi} g(x)x\mathrm{d}\theta\mathrm{d}x$ equals $E(M + P)$ and is convergent, $f'(d)$ is also convergent.

APPENDIX D: PROOF OF THEOREM 3

Regarding r as a variable, we substitute the notations as follows: $S \to S(r)$, $f(d) \to f(r, d)$, $f'(d) \to \partial f(r, d)/\partial d$, $\mathrm{CRLB}(d) \to \mathrm{CRLB}(r, d)$, $g(d) \to g(r, d)$. According to (14), we have

$$g(r, dr) = \int_{k\log d}^{\infty} \frac{\mathrm{e}^{-\frac{z^2}{2\sigma_{\mathrm{dB}}^2}}}{\sqrt{2\pi}\sigma_{\mathrm{dB}}}\mathrm{d}z = g(1, d)$$

By (13), we have

$$\begin{aligned} f(r,dr) &= \int_0^\infty \int_0^{2\pi} g(r,x) \\ &\quad \times g(r,\sqrt{x^2+(dr)^2-2xdr\cos\theta})x\mathrm{d}\theta\mathrm{d}x \\ &= \int_0^\infty \int_0^{2\pi} g(r,xr) \\ &\quad \times g(r,r\sqrt{x^2+d^2-2xd\cos\theta})r^2x\mathrm{d}\theta\mathrm{d}x \\ &= r^2\int_0^\infty \int_0^{2\pi} g(1,x) \\ &\quad \times g(1,\sqrt{x^2+d^2-2xd\cos\theta})x\mathrm{d}\theta\mathrm{d}x \\ &= r^2 f(1,d) \end{aligned}$$

Moreover, we can obtain

$$\left.\frac{\partial f(x,y)}{\partial y}\right|_{x=1,y=d} = \frac{1}{r}\times \left.\frac{\partial f(x,y)}{\partial y}\right|_{x=r,y=dr} \quad (29)$$

By $S(r) = r^2 S(1)$ (based on (15)), (23) and the above formulas, we can obtain

$$\begin{aligned} &\mathrm{CRLB}(r,dr) = \mathrm{CRLB}(1,d) \text{ equivalently,} \\ &\mathrm{CRLB}(r,d) = \mathrm{CRLB}\left(1,\frac{d}{r}\right) \end{aligned} \quad (30)$$

ACKNOWLEDGEMENTS

The work was supported by ARC (Australian Research Council) under DP-110100538 and NICTA. C. Yu is an ARC Queen Elizabeth II Fellow and is also supported by the Overseas Expert Program of Shandong Province. B. Huang was with the Australian National University when the work was performed, and he was also supported by ARC, NICTA and the Key Laboratory of Computer Networks of Shandong Province. NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. This material is based on research sponsored by the Air Force Research Laboratory, under agreement number FA2386-09-1-4136. The US Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the US Government.

REFERENCES

1. Mao G, Fidan B, Anderson BDO. Wireless sensor network localization techniques. *Computer Networks* 2007; **51**(10): 2529–2553.
2. Li L, Halpern JY, Bahl P, Wang Y, Wattenhofer R. A cone-based distributed topology-control algorithm for wireless multi-hop networks. *IEEE/ACM Transactions on Networking* 2002; **13**: 147–159.
3. Siripongwutikorn P, Thipakorn B. Mobility-aware topology control in mobile ad hoc networks. *Computer Communications* 2008; **31**(14): 3521–3532.
4. Sitanayah L, Datta A, Cardell-Oliver R. Heuristic algorithm for finding boundary cycles in location-free low density wireless sensor networks. *Computer Networks* 2010; **54**(10): 1630–1645.
5. Bulusu N, Heidemann J, Estrin D. GPS-less low cost outdoor localization for very small devices. *IEEE Personal Communications Magazine* 2000; **7**(5): 28–34.
6. Niculescu D, Nath B. Ad hoc positioning system (APS), In *Proc. of the IEEE Globecom*, San Antonio, TX, USA, 2001; volume 5, 2926–2931.
7. Nagpal R, Shrobe H, Bachrach J. Organizing a global coordinate system from local information on an ad hoc sensor network, In *Proc. of the ACM/IEEE IPSN*, Palo Alto, CA, USA, 2003; 333–348.
8. Shang Y, Ruml W, Zhang Y, Fromherz MPJ. Localization from mere connectivity, In *Proc. of the ACM MobiHoc*, Annapolis, Maryland, USA, 2003; 201–212.
9. Buschmann C, Pfisterer D, Fischer S. Estimating distances using neighborhood intersection, In *Proc. of the IEEE Emerging Technologies and Factory Automation*, Prague, Czech, 2006; 314–321.
10. Buschmann C, Hellbrück H, Fischer S, Kröller A, Fekete S. Radio propagation-aware distance estimation based on neighborhood comparison, In *Proc. of the 4th European Conference on Wireless Sensor Networks*, Delft, The Netherlands, 2007; 325–340.
11. Rappaport T. *Wireless Communications: Principles and Practice*. Prentice Hall PTR: New Jersey, USA, 2001.
12. Meester R, Roy R. *Continuum Percolation*. Cambridge University Press: Cambridge, UK, 1996.
13. Franceschetti M, Meester R. *Random Networks for Communication: From Statistical Physics to Information Systems*. Cambridge University Press, 2007.
14. Ta X, Mao G, Anderson BDO. On the Connectivity of Wireless Multi-hop Networks with Arbitrary Wireless Channel Models. *IEEE Communications Letters* 2009; **13**(3): 181–183.

15. Ramaswami R, Parhi KK. Distributed scheduling of broadcasts in a radio network, In *Proc. of the IEEE INFOCOM*, Ottawa, Ont., Canada, 1989; 497–504 vol.2.
16. Mukherjee S, Avidor D. Connectivity and transmit-energy considerations between any pair of nodes in a wireless ad hoc network subject to fading. *IEEE Transactions on Vehicular Technology* 2008; **57**(2): 1226–1242.
17. Orriss J, Barton SK. Probability distributions for the number of radio transceivers which can communicate with one another. *IEEE Transactions on Communications* 2003; **51**(4): 676–681.
18. Bettstetter C, Hartmann C. Connectivity of wireless multihop networks in a shadow fading environment. *Wireless Networks* 2005; **11**(5): 571–579.
19. Miorandi D, Altman E. Coverage and connectivity of ad hoc networks in presence of channel randomness, In *Proc. of the IEEE INFOCOM*, Miami, FL, USA, 2005; 491–502.
20. Miorandi D. The impact of channel randomness on coverage and connectivity of ad hoc and sensor networks. *IEEE Transactions on Wireless Communications* 2008; **7**(3): 1062–1072.
21. Mao G, Anderson BDO, Fidan B. Path loss exponent estimation for wireless sensor network localization. *Comput. Netw.* 2007; **51**: 2467–2483.
22. Chitte SD, Dasgupta S, Ding Z. Distance estimation from received signal strength under log-normal shadowing: bias and variance. *IEEE Signal Processing Letters* 2009; **16**(3): 216–218.
23. Akyildiz IF, Su W, Sankarasubramaniam Y, Cyirci E. Wireless sensor networks: a survey. *Computer Networks* 2002; **38**(4): 393–422.
24. Pan J, Hou YT, Cai L, Shi Y, Shen SX. Topology control for wireless sensor networks, In *Proc. of the ACM MobiCom*, San Diego, CA, USA, 2003; 286–299.
25. Chen B, Jamieson K, Balakrishnan H, Morris R. Span: an energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. *Wireless Networks* 2002; **8**(5): 481–494.
26. Wang X, Xing G, Zhang Y, Lu C, Pless R, Gill C. Integrated coverage and connectivity configuration in wireless sensor networks, In *Proc. of the ACM SenSys*, Los Angeles, CA, USA, 2003; 28–39.
27. Shang Y, Shi H. A new density control algorithm for WSNs, In *Proc. of the IEEE LCN*, Tampa, FL, USA, 2004; 577–578.
28. Ye F, Zhong G, Cheng J, Lu S, Zhang L. PEAS: A robust energy conserving protocol for long-lived sensor networks, In *Proc. of the IEEE ICDCS*, Providence, RI, USA, 2003; 28–37.
29. Yen L, Cheng Y. Range-based sleep scheduling (RBSS) for wireless sensor networks. *Wireless Personal Communications* 2009; **48**(3): 411–423.
30. Wireless sensor network localization measurement repository. http://www.eecs.umich.edu/hero/localize/, 2011.
31. Ma D, Meng JE, Wang B, Lim HB. A novel approach towards source-to-destination distance estimation in wireless sensor networks, In *Proc. of the Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, Melbourne, Australia, 2009; 463–467.
32. Lederer S, Wang Y, Gao J. Connectivity-based localization of large scale sensor networks with complex shape, In *Proc. of the IEEE INFOCOM*, Phoenix, AZ, USA, 2008; volume 5, 789–797.
33. Lim H, Hou JC. Distributed localization for anisotropic sensor networks. *ACM Transactions on Sensor Networks* 2009; **5**: 1–26.
34. Wang Y, Li K, Wu J. Distance estimation by constructing the virtual ruler in anisotropic sensor networks, In *Proc. of the IEEE INFOCOM*, San Diego, CA, USA, 2010; 1172–1180.

AUTHORS' BIOGRAPHIES

Baoqi Huang received his BEng degree from Inner Mongolia University, China, and his MS degree from Peking University, China, both in computer science. Between March 2008 and September 2011, he was pursuing his PhD degree in information engineering from the Australian National University, Canberra, Australia. Currently, he is with the school of computer, Inner Mongolia University, China. His research interests include wireless sensor networks and mobile *ad hoc* networks.

Changbin (Brad) Yu received his BEng degree with first class honors in computer engineering from Nanyang Technological University, Singapore, in 2004 and his PhD in information engineering from the Australian National University, Canberra, Australia, in 2008. He is now a an ARC Queen Elizabeth II Fellow at the Australian National University and adjunct at NICTA Ltd and Shandong Computer Science Center. He was a recipient of an ARC Australian Postdoctoral Fellowship in 2008, the Chinese Government Award for Outstanding Chinese Students Abroad in 2006, the Australian Government's Endeavour

Asia Award in 2005, and an ARC Queen Elizabeth II Fellowship in 2011. His current research interests include control of autonomous formations, multi-agent systems, sensor networks, and graph theory.

Brian D. O. Anderson is a distinguished professor at the Australian National University and a distinguished researcher in National ICT Australia. He obtained his PhD in electrical engineering from Stanford University in 1966. He is a fellow of the IEEE, IFAC, the Australian Academy of Science, the Australian Academy of Technological Sciences and Engineering, and the Royal Society and a foreign associate of the National Academy of Engineering. His current research interests are in distributed control, including control of UAV formations, sensor networks, and econometric modeling.

Guoqiang Mao received his bachelor's degree in electrical engineering, his master's degree in engineering, and his PhD degree in telecommunications engineering in 1995, 1998, and 2002, respectively. In December 2002, he joined the School of Electrical and Information Engineering, the University of Sydney, Australia, where he is a senior lecturer now. He has published more than 50 papers in prestigious journals and refereed conference proceedings. He was listed in the 25th Anniversary Edition of Marquis 'Who's Who in the World' (2008) and in the 9th (2007) and 10th (2008) Anniversary Editions of Marquis 'Who's Who in Science and Engineering' for 'exceptional achievements in Science and Engineering'. His research interests include wireless localization techniques, wireless multihop networks, graph theory, and its application in networking, telecommunications traffic measurement, analysis and modeling, and network performance analysis.

2011 9th IEEE International Conference on
Control and Automation (ICCA)
Santiago, Chile, December 19-21, 2011

A Generic Bias-Correction Method with Application to Scan-Based Localization

Yiming Ji, Changbin Yu Brian D.O. Anderson and Samuel P. Drake

***Abstract*—In previous work a method was proposed to determine the bias in localization algorithms using range or bearing data. In this paper the method is extended to be more generic; in particular, different types of measurement data are permitted, and there may be more measurements than there are variables to estimate. The method combines the Taylor series and Jacobian matrices to determine the bias, and leads to an easily calculated analytical bias expression, despite the general unavailability of analytic expressions for the solution of most localization problems. The method is used to estimate the bias in scan-based localization. Monte Carlo simulation results verify the performance of the proposed method in this context.**

Keywords: **Bias; Taylor series; Scan-based localization; Geolocation; Passive Localization; Targeting; Tracking**

I. INTRODUCTION

Bias is a term in estimation theory and is defined as the difference between the expected value of a parameter estimate and its true value [1]. If the bias of a particular estimation scheme can be estimated then it can be removed. Of course, in a particular single instance of the estimation problem, the correction may worsen the quality of the estimate; but over a number of experiments, it may be expected to improve the quality of the estimate, by lowering the mean square error.

A common class of estimation problems are those localization estimates, i.e. determination of the position of a target, e.g. a sound emitter, or an electromagnetic wave emitter, using some form of measurements, such as range, time difference of arrival, bearing, etc. In almost all practical localization situations, measurement errors are inevitable, and these lead to errors in estimating the true target position. In [2] Picard et al. discussed several models for estimating the bias in the range measurements, and presented a set of iterative algorithms that minimize the bias and provide maximum likelihood position estimates. In [3], Gavish et al. presented analytical expressions of bias which permit performance comparison for two well known bearing-only location techniques, viz. the maximum likelihood and the Stansfield estimators. Further in [4], Drake et al. presented an introduction to tensor algebra with some application examples in estimation theory. One of the tensor algebra applications proposed in the paper treats the bias in estimating a nonlinear function of a variable of which one has an observation contaminated by zero mean additive noise. The method involves expanding the nonlinear function around the noiseless observation value using a Taylor series which is truncated at the second order. The expected value of the first order term is zero and the expected value of the second order term introduces a bias. As an example the localization of a target is considered using noisy range and bearing information from a monostatic radar; with independent and gaussian zero mean noises contaminating each measurement, there is a systematic bias in the target estimate, such that the estimated position on average is closer to the radar than the true target position.

In our previous work [5, 6, 7], we proposed a method to determine and correct the bias in localization algorithms. We hypothesized the existence of a localization mapping $\mathbf{g}$ (which maps the vector of measurements to a position vector estimate). Following the lead of [4], we viewed this mapping using a Taylor series expansion around the nominal noiseless measurements that in principle are associated with the correct target position. The expansion is to second-order in the measurement noise as the first order term has zero mean and the expected values of the second-order term as expressible in terms of the derivatives of $\mathbf{g}$. However, in a localization problem it may actually be very hard to calculate the derivatives of $\mathbf{g}$ analytically. In contrast, the inverse mapping of $\mathbf{g}$ (call it $\mathbf{f}$) which maps the target position to a (noiseless) sensor measurement can often be obtained analytically, together with its derivatives. Therefore, we introduce the Jacobian matrix of $\mathbf{f}$ to compute the derivatives of the localization mapping $\mathbf{g}$ in terms of the derivatives of $\mathbf{f}$, resulting in a simple calculation of bias. In comparison with the approach presented in [3], the Monte Carlo simulation results demonstrate a clearly better performance for our bias correction method [7].

In our previous work however, the analysis of the proposed method was restricted to localization problems using only range or bearing measurements. In this paper we present the bias correction method in a more generic way allowing an arbitrary number of noisy measurements which are not restricted to being either range or bearing. To demonstrate the performance of the proposed method, the generic bias correction method is applied by way of example to scan-based localization algorithms, which have not been analyzed in our previous work. In the process of applying the proposed method to scan-based localization, the original bias correction method needs to be adjusted in a minor way to allow for certain correlations in the measurement errors.

The rest of this paper is organized as follows. In Section II the background and motivation of our work are presented. The generic bias correction method is proposed in Section III. In Section IV, we apply the bias correction method to the scan-based localization problem. Monte Carlo simulation results are also provided in Section IV. Section V summarizes the main results of the paper.

II. BACKGROUND AND MOTIVATION

A. Background

A brief review of estimation bias will be presented in this section.

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ and $\mathbf{\Theta} = (\theta_1, \theta_2, \ldots, \theta_N)^T$ denote the target position and noiseless measurement vector respectively. Let $\mathbf{f}: \mathbf{x} \rightarrow \mathbf{\Theta}$ denote the associated mapping, which is almost always analytically computable. Let $\mathbf{g}: \mathbf{\Theta} \rightarrow \mathbf{x}$ denote the inverse localization mapping; thus with $\mathbf{\Theta} =$

Y. Ji and B. Anderson are with National ICT Ltd. Australia-NICTA and Research School of Engineering, The Australian National University Canberra ACT 2601, Australia. Email: {yiming.ji, brian.anderson}@anu.edu.au

C. Yu is with Research School of Engineering, The Australian National University Canberra ACT 2601, Australia. Email: brad.yu@anu.edu.au

S. P. Drake is with Electronic Warfare and Radar Division, Defence Science and Technology Organization (DSTO) Edinburgh, SA 5111, Australia. Email: sam.drake@dsto.defence.gov.au

978-1-4577-1474-0/11/$26.00 ©2011 IEEE

$\mathbf{f}(\mathbf{x})$, there holds $\mathbf{x} = \mathbf{g}(\mathbf{\Theta})$. In order that this mapping can be well defined, it is necessary that $N \geq n$. In case $N > n$, usually the set of equations $\mathbf{f}(\mathbf{x}) = \mathbf{\Theta}$ yielding $\mathbf{x}$ from $\mathbf{\Theta}$ will be overdetermined, while in case $N = n$, there may be two or more solutions (but for generic geometries only a finite number); in this case, the selection of the correct solution requires some additional information regarding the target position.

In practice, noise in the measurements is inevitable. We denote this noise by $\delta\mathbf{\Theta} = (\delta\theta_1, \delta\theta_2, ..., \delta\theta_N)^T$. The $\delta\theta_i$ are generally assumed to be independent Gaussian random variables with zero mean and known variances σ_i^2, which may be the same. The error in the target position resulting from an estimation procedure using noisy data is denoted as $\delta\mathbf{x}$. If we define $\tilde{\mathbf{\Theta}} = \mathbf{\Theta} + \delta\mathbf{\Theta}$ and $\tilde{\mathbf{x}} = \mathbf{x} + \delta\mathbf{x}$ the localization amounts to solving $\mathbf{f}(\tilde{\mathbf{x}}) = \tilde{\Theta}$ for $\tilde{\mathbf{x}}$. If the function $\mathbf{g}$ is known, then in effect we are implementing the following equation

$$\tilde{\mathbf{x}} = \mathbf{x} + \delta\mathbf{x} = \mathbf{g}(\Theta + \delta\Theta) = \mathbf{g}(\tilde{\Theta}) \tag{1}$$

If $\mathbf{g}$ is a non-linear function then this process will lead to bias in the target position estimate [8]. Suppose, by way of a thought experiment, that in estimating the value of x the measurement process equation (1) was repeated M times. As $M \to \infty$, the average of the estimates would go to :

$$E[\tilde{x}_i] = E[g_i(\tilde{\Theta})] \tag{2}$$

Now note that if g_i is nonlinear we would have:

$$E[\tilde{x}_i] = E[g_i(\tilde{\Theta})] \neq g_i(E[\tilde{\Theta}]) = g_i(\Theta) = x_i$$

The bias in our estimated of x_i is defined as the difference between the expected value of x_i and the true value of x_i, i.e.

$$Bias_{x_i} = E[\tilde{x}_i] - x_i = E[g(\tilde{x}_i)] - x_i$$

If computable, the bias can be used to systematically correct any single estimate from any single measurement set.

B. Motivation

From the above analysis, we can see that if the estimation mapping $\mathbf{g}$ is nonlinear and the sensor measurements are noisy, bias is present. In practice, these two conditions are presented in most localization scenarios. Since bias is a systematic and possibly computable error, it is desirable to remove it.

In [4], Drake et al. gave a short introduction to tensor algebra and provided a few sample applications. One such application was concerned with bias arising in non-linear estimation problems with noisy measurements. To determine the bias they considered, as we do, $\tilde{x}_i = g_i(\tilde{\theta})$. Assuming the estimator mapping $\mathbf{g}$ is well defined, they expanded the function g_i by a Taylor series and truncated it at second order:

$$\begin{aligned} x_i + \delta x_i &= g_i(\tilde{\theta}_1, \tilde{\theta}_2, ..., \tilde{\theta}_N) \\ &= g_i(\theta_1 + \delta\theta_1, \theta_2 + \delta\theta_2, ..., \theta_N + \delta\theta_N) \\ &\approx g_i(\theta_1, \theta_2, ..., \theta_N) + \sum_{j=1}^{N} \frac{\partial g_i}{\partial \theta_j}\delta\theta_j \\ &\quad + \frac{1}{2!}\sum_{j=1}^{N}\sum_{l=1}^{N} \delta\theta_j\delta\theta_l \frac{\partial^2 g_i}{\partial\theta_j\partial\theta_l} \end{aligned}$$

Assuming as is often the case that the measurement errors are independent Gaussian random variables with zero mean and known variance (the variance of measurement errors would have to be obtained from manufacturer or experimental data), the approximate bias expression is:

$$E(\delta x_i) = \frac{1}{2!}\sum_{j=1}^{N} \sigma_j^2 \frac{\partial^2 g_i}{\partial\theta_j^2} \tag{3}$$

For some estimators, the mapping $\mathbf{g}$ can be obtained analytically. However in some situations including many localization problems, finding the analytic $\mathbf{g}$ analytically is very hard or even impossible. If $\mathbf{g}$ can not be obtained then equation (3) can not be evaluated, and we need a new method to analytically obtain the derivatives. The bias in our estimated of x_i is defined as the difference between the expected value of x_i and the true value of x_i.

The key to do this is to notice that $\mathbf{g}$ is the inverse of the mapping $\mathbf{f}$ for which often an analytic form is known. Below we show how to use the mapping $\mathbf{f}$ and its derivatives to calculate the derivatives of $\mathbf{g}$ using the Jacobian identity, ultimately resulting in an estimate of the bias.

III. Bias Correction Method

To begin, we first assume $N = n$ (the number of sensor measurements N is equal to the dimension of position coordinates of the target n). We assume further that $\mathbf{f}$ is a known analytic function. Because $\mathbf{f}$ and $\mathbf{g}$ are inverse mappings, the Jacobian identity holds:

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \frac{\partial f_N}{\partial x_1} & \cdots & \frac{\partial f_N}{\partial x_n} \end{bmatrix} \begin{bmatrix} \frac{\partial g_1}{\partial \theta_1} & \cdots & \frac{\partial g_1}{\partial \theta_N} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \frac{\partial g_n}{\partial \theta_1} & \cdots & \frac{\partial g_n}{\partial \theta_N} \end{bmatrix} = \mathbf{I}_n \tag{4}$$

By rearranging the equation set (4) we can obtain analytical expressions for the $\frac{\partial g_i}{\partial \theta_j}$ $(i = 1, 2, ..., n; j = 1, 2, ..., N)$ in terms of $\frac{\partial f_i}{\partial x_j}$, and thus as analytic functions of the x_i. For ease of exposition we use g_j^i to denote the expressions of $\frac{\partial g_i}{\partial \theta_j}$ as functions of $x_1, x_2, ..., x_n$. To obtain second derivatives, let us use $\frac{\partial g_1}{\partial \theta_1}$ as an example to illustrate the general approach. Starting with

$$\frac{\partial g_1}{\partial \theta_1} = g_1^1 \tag{5}$$

we differentiate with respect to x_1 first, and thereby obtain $\frac{\partial g_1}{\partial {\theta_1}^2}\frac{\partial f_1}{\partial x_1} + ... + \frac{\partial g_1}{\partial \theta_1 \partial \theta_i}\frac{\partial f_i}{\partial x_1} + ... + \frac{\partial g_1}{\partial \theta_1 \partial \theta_N}\frac{\partial f_N}{\partial x_1} = \frac{\partial g_1^1}{\partial x_1}$. If we further differentiate the equation (5) with respect to $x_2, ..., x_n$ we can obtain an equation set as follows:

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_N}{\partial x_1} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \frac{\partial f_1}{\partial x_n} & \cdots & \frac{\partial f_N}{\partial x_n} \end{bmatrix} \begin{bmatrix} \frac{\partial^2 g_1}{\partial {\theta_1}^2} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial^2 g_1}{\partial \theta_1 \partial \theta_N} \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1^1}{\partial x_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial g_1^1}{\partial x_n} \end{bmatrix} \tag{6}$$

Note that the quantities on the right side of this equation are all expressible analytically as functions of $x_1, x_2, ..., x_n$. Likewise the entries $\frac{\partial f_i}{\partial x_j}$ in the matrix on the left are known functions of $x_1, x_2, ...x_n$. Hence by solving the equation set (6), we can obtain a formula for $\frac{\partial^2 g_1}{\partial \theta_1^2}$ as a function of $x_1, x_2, ..., x_n$. The formulas for $\frac{\partial^2 g_i}{\partial \theta_j^2}$ for all i, j can be obtained in the same way. Substituting the formulas

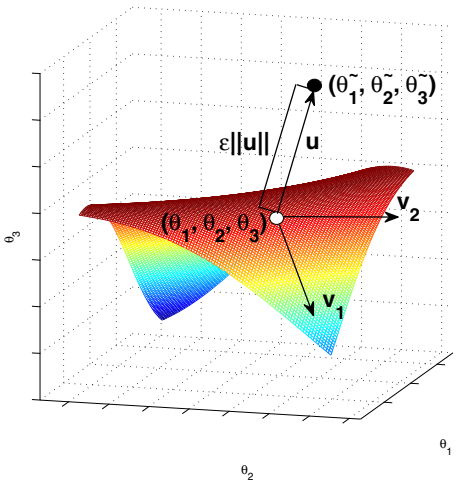


Fig. 1. Introduce one extra variable (Here N=3 and n=2). The surface is the set of points $(\theta_1, \theta_2, \theta_3) = (f_1(x,y), f_2(x,y), f_3(x,y))$ obtained as x, y vary.

into equation (3) we can finally obtain the easily-calculated expressions for the bias in terms of f_i and its derivatives.

The calculation is used as follows. We first obtain the estimated position of the target by using an existing localization algorithm. Then we can input the estimated target location into the obtained analytical expression for the bias. Finally we can improve (on a statistical basis at least) the accuracy of the localization by subtracting the obtained bias, viz. $\tilde{x_i} - bias_{x_i} (i = 1, 2, ..., n)$. (The process could in fact be iterated, but the effect of even one more iteration on the corrected position estimate will almost always be marginal).

If the number of measurements N is greater than the number of scalar position variables n, we cannot obtain the equations (4) and (6), and so we cannot straightforwardly express the bias using the derivatives of $\mathbf{f}$. Indeed, while the noise-free equation $\mathbf{f}(\mathbf{x}) = \Theta$ is overdetermined, the noisy equation $\mathbf{f}(\tilde{\mathbf{x}}) = \tilde{\Theta}$ in general will have no solution. The localization problem is typically solved by something like a least squares approach[1], and for the purposes of bias determination, we build on this approach too, to introduce $N - n$ extra variables into $\mathbf{f}$ thereby making $n = N$.

For ease of exposition, here we take $N = n + 1$, which means just one extra variable needs to be introduced. Consider N-dimensional space, with axes corresponding to the N measurements. Assume an $(N-1)$-dimensional hypersurface (illustrated in Figure 1 for the case $N = 3$) consists of points which correspond to all sets of noiseless measurements $(\theta_1, \theta_2, ..., \theta_N)$, i.e. $\theta_i = f_i(x_1, x_2, ...x_n)$ for $i = 1, 2, ..., N$. According to the least squares method, we can consider choosing $x_1, x_2, ..., x_n$ to minimize the following cost function:

$$F_{\text{cost-function}}(\mathbf{x}, \tilde{\Theta}) = \sum_{i=1}^{N} (f_i - \tilde{\theta}_i)^2 = \sum_{i=1}^{N} \delta\theta_i^2 \qquad (7)$$

In fact, the least squares method attempts to find a point $(\theta_1, \theta_2, ..., \theta_N)$ (the white point in Figure 1) on the surface corresponding to an obtained set of noisy measurements $(\tilde{\theta}_1, \tilde{\theta}_2, ..., \tilde{\theta}_N)$ (the black point in Figure 1 which is generically off the surface) to minimize the distance between the two points. Hence the white point must be the orthogonal projection of the black one onto the surface, or the black point must be on the normal vector to a tangent plane of the surface passing through the white one.

The n tangent vectors at any point on the n dimensional surface are given by vectors as follows:

$$\mathbf{v}_i = [\frac{\partial f_1}{\partial x_i}, \frac{\partial f_2}{\partial x_i}, ..., \frac{\partial f_N}{\partial x_i}]^T \quad i = 1, 2, ..., n \qquad (8)$$

The normal vector to the surface ($\mathbf{u}$) is formed from the cross product of the tangent vectors ($\mathbf{v}_i$), that is: $\mathbf{u}$ [9] of the surface through the white point:

$$\mathbf{u} = [u_1, u_2, ..., u_N]^T = \mathbf{v}_1 \times \mathbf{v}_2 ... \times \mathbf{v}_n \qquad (9)$$

Note that in the noiseless case $\Theta = \mathbf{f}(x_1, x_2, ...x_n)$ where $\mathbf{f}$ can be written down easily according to the geometry of the sensors and target. The black point can be regarded as an image obtained with a new analytical mapping $\mathbf{F} = (F_1, F_2, ..., F_N)^T : R^N \rightarrow R^N$ corresponding to moving from $\mathbf{f}$, which is a known function of $x_1, x_2, ...x_n$ along the normal vector for a distance $\varepsilon||\mathbf{u}||$. The new mapping $\mathbf{F}$ is a known analytic function of $x_1, x_2, ...x_n$ and ε:

$$\tilde{\Theta} = \mathbf{F}(\tilde{\mathbf{x}}, \varepsilon) = \mathbf{f}(\tilde{\mathbf{x}}) + \varepsilon\mathbf{u} \qquad (10)$$

(Of course, $\mathbf{u}$ is an analytic function of $x_1, x_2, ...x_n$).

Introduction of the extra variable ϵ means that equation (10) is not an overdetermined equation set, and $\mathbf{F}$ is in principle invertible. Therefore we can consider the new localization mapping (call it $\mathbf{G}$) as the inverse mapping of $\mathbf{F}$. We can then proceed along the same lines as previously to determine the bias.

When the number of input measurements N exceeds $n+1$, the situation is similar to the case $N = n + 1$. More details can be obtained in [6, 7].

IV. Application to the Scan-Based Localization Problem

In order to demonstrate the effectiveness of the proposed bias-correction method, we apply the proposed method to the scan-based localization problem, which is studied in [10].

A. Review of Scan-Based Localization

The target is assumed to be an emitter with a (generally mechanically) rotating radar antenna with a narrow beam; the scan direction and scan rate are assumed constant and indeed known to each receiver, which records the time instants at which the rotating beam passes the receiver [10]. For ease of exposition, here we consider a situation with one emitter and three receiving sensors, and assume that all lie in a plane. Figure 2 shows an emitter scanning across three receivers (i.e., receiver 1, receiver 2 and receiver 3) at the times t_1, t_2 and t_3. The emitter is scanning clockwise. For localization, the separate time values are not important, but rather their differences, $t_{12} = t_2 - t_1$ and $t_{23} = t_3 - t_2$, are. In fact, we treat t_{12}, t_{23} as quasi-measurements. Each quasi-measurement (together with knowledge of the scan rate and direction) in the noiseless case defines a circle of computable center and radius, and indeed an arc of such a circle on which the emitter must lie. The pair of sensors determining the time difference lie at the end points of the arc. The intersection of two such circular arcs defines the emitter position; thus there is a vector function $\mathbf{g} = (g_1, g_2)^T$, the localization mapping, of the two variables t_{12}, t_{23}, with $\mathbf{g}$ embodying a formula for the intersection of two circular arcs. Finding an analytic expression for the mapping is a significant challenge.

[1]The least squares approach is equivalent to a maximum likelihood approach when all noise variances are the same (and measurement noises are independent zero mean Gaussian random variables). Weighted least squares can capture variations on this.

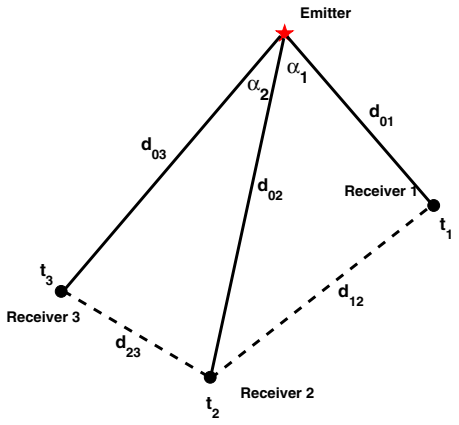


Fig. 2. An emitter scanning across three receivers; The red star indicates the radar's location whereas the filled circles indicates the location of the receivers.

The two factors causing bias to arise in an estimate, viz. nonlinearity of the estimation mapping and noise in the measurements, are present in any practical scan-based localization problem. Accordingly, we seek to apply our method to use the inverse of **g** (viz. the mapping **f** from target position to measurements) to obtain a formula for the bias.

B. Analytic expression for the mapping from emitter position to quasi-measurements

Now we aim to obtain an analytic form for the function **f** which maps the emitter position (x, y) to the quasi-measurements (t_{12}, t_{23}). Let α_1, α_2 denote the angles subtended at the emitter by the lines joining it to the two pairs of physical sensors, see Figure 2. Since the scan rate ω is a known constant, we can obtain the following equations:

$$\alpha_i = \omega t_{i,i+1} \qquad i = 1, 2 \tag{11}$$

Given the three receivers at known locations $p_1 = (x_1, y_1)$, $p_2 = (x_2, y_2)$ and $p_3 = (x_3, y_3)$, it is straightforward to see that

$$\alpha_i = \arccos \frac{d_{0,i}^2 + d_{0,i+1}^2 - d_{i,i+1}^2}{2 d_{0,i} d_{0,i+1}} \qquad i = 1, 2 \tag{12}$$

where

$$d_{0,i} = \sqrt{(x - x_i)^2 + (y - y_i)^2}, \quad i = 1, 2, 3$$

$$d_{i,i+1} = \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}, \quad i = 1, 2$$

Substituting equations (12) into (11), we can obtain the following formulas:

$$t_{i,i+1} = f_i(x, y) = \arccos \frac{d_{0,i}^2 + d_{0,i+1}^2 - d_{i,i+1}^2}{2 d_{0,i} d_{0,i+1} \omega} \qquad i = 1, 2 \tag{13}$$

These last equations provide an analytic formula for $\mathbf{f} = (f_1, f_2)$, which is the inverse mapping of localization process $\mathbf{g} = (g_1, g_2)$.

When the number of receivers is larger than 3, we can use the least-squares based method proposed in Section III to introduce extra variables.

C. Bias-Correction in Scan-Based Localization

As the first step in obtaining an expression for the bias, we Taylor expand the localization mappings g_1 and g_2 to the second order terms, as described in Section III. However, in the scan-based localization problem, equation (3) requires adjustment. To see why, note that noise in the time-of-arrival (TOA) measurements can be modelled as follows:

$$\tilde{t_i} = t_i + \delta t_i \quad i = 1, 2, 3 \tag{14}$$

where the δt_i are assumed to be i.i.d Gaussian random variables with zero mean and known variance σ^2.

However in scan-based localization, the physical measurements are replaced by quasi-measurements $\tilde{t}_{12} = \tilde{t}_2 - \tilde{t}_1$ and $\tilde{t}_{23} = \tilde{t}_3 - \tilde{t}_2$. This leads to

$$\tilde{t}_{12} = \tilde{t}_2 - \tilde{t}_1 = t_{12} + \delta t_{12} \tag{15}$$

$$\tilde{t}_{23} = \tilde{t}_3 - \tilde{t}_2 = t_{23} + \delta t_{23} \tag{16}$$

where δt_{12} and δt_{23} are no longer independent and have covariance matrix Σ given by

$$\Sigma = 2\sigma^2 \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix} \tag{17}$$

where $2\sigma^2$ is the variance of an individual time-difference measurement. Note that the means of δt_{12} and δt_{23} remain zero.

Now the approximate bias expression for three receivers is as follows:

$$E(\delta x) = \frac{1}{2!}[2\sigma^2 \frac{\partial^2 g_1}{\partial t_{12}^2} - 2\sigma^2 \frac{\partial^2 g_1}{\partial t_{12} \partial t_{23}} + 2\sigma^2 \frac{\partial^2 g_1}{\partial t_{23}^2}] \tag{18}$$

$E(\delta y)$ can be obtained in the same way. The remaining calculations are the same as those described in Section III.

D. Simulation Results

In this subsection, the simulation results will be shown to demonstrate the performance of the proposed bias-correction method in scan-based localization problems. All the simulated data is provided by the Defence Science Technology Organization (DSTO).

The simulations were done using DSTO's synthetic integration lab (SIL). The still accurately simulates existing radar and receiver systems, it includes accurate, senor, emitter, terrain and propagation models. The fidelity of the SIL means that there is no significant difference between its simulation results and those generated by field tests. Simulation results are provided in 2-dimensional space with two scenarios: (1) three receivers and one emitter (2) four receivers and one emitter. Different emitter positions are considered.

The simulation set-up is as follows:

- The measurement error δt_i for each sensor is produced by an independent Gaussian distribution with zero mean and known variance σ^2. The level of noise (the standard deviation σ) is adjusted in the simulation set as 0.05, 0.1 or 0.15 seconds for TOA measurements t_i.
- All the simulation results are obtained from 1000 Monte Carlo experiments.
- In the simulations the bias is considered as the average absolute distance (average of 1000 experimental results) between the true emitter position and the estimated emitter position. In the simulation figures it is termed the 'average absolute distance error'.[2]

[2]In practice, the bias is a vector whose entries can be negative or positive. Here we only focus on how large the bias is. Therefore the absolute distance between the estimated target position and the true position is used to evaluate the bias.

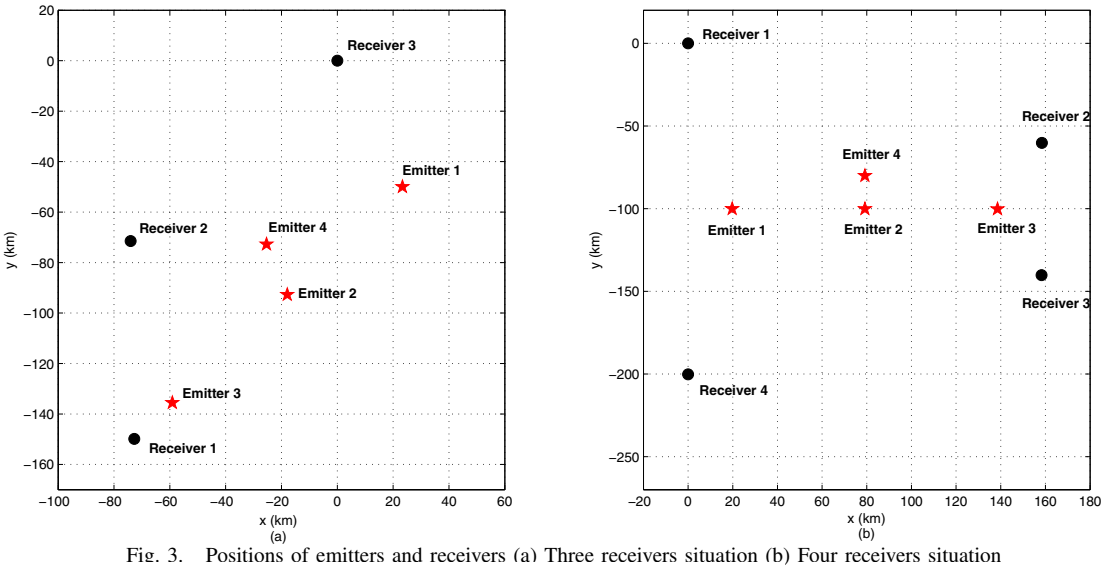


Fig. 3. Positions of emitters and receivers (a) Three receivers situation (b) Four receivers situation

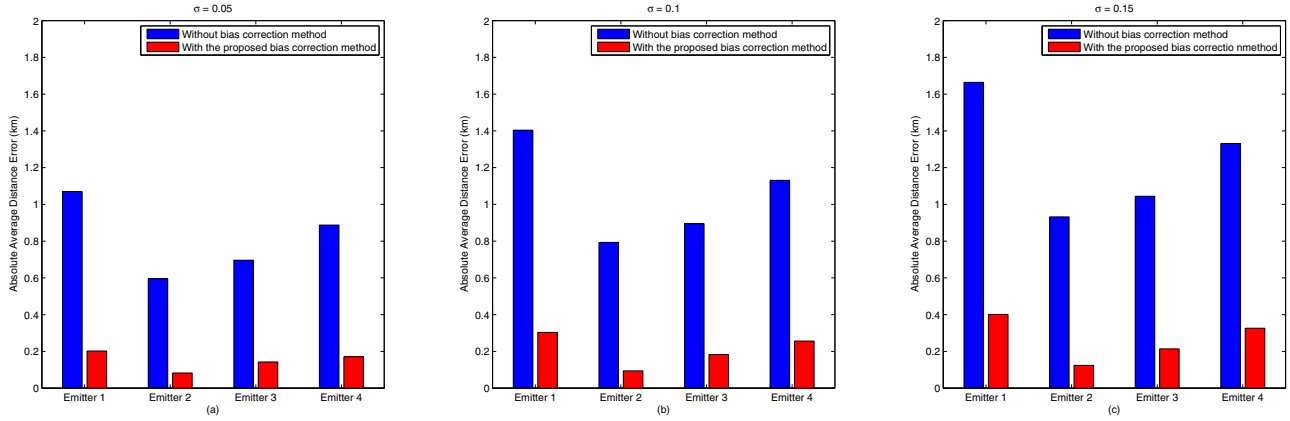


Fig. 4. Three receivers scenario with different noise level (a) σ=0.05 (b) σ=0.10 (c) σ=0.15

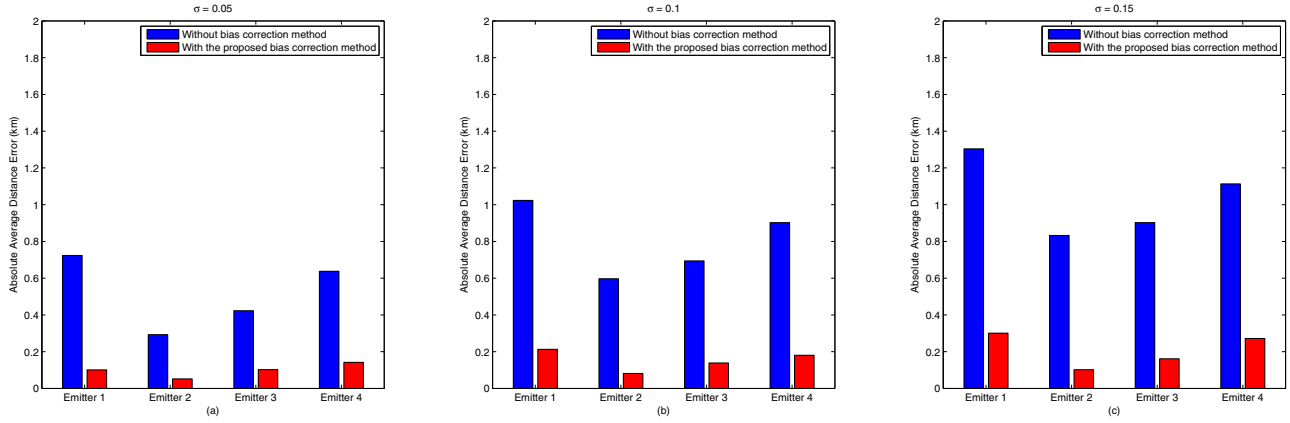


Fig. 5. Four receivers scenario with different noise level (a) σ=0.05 (b) σ=0.10 (c) σ=0.15

- 'Analytical bias' denotes the bias obtained by using the analytical expression derived from the proposed bias-correction method.
- 'Experimental bias' denotes the bias obtained by using simulation.
- The scan rate is $\frac{4\pi}{5}$ radians per second. Thus one period is 2.5 seconds, and so the standard deviation of TOA measurements is between 2% and 6% of a period.
- The distance unit used in the simulations is kilometers (km).
- The time unit and the angle unit used in the simulations are seconds and radians.

1) Three Receivers Scenario

In this situation, three receivers give rise to two quasi-measurements (t_{12} and t_{23}). Therefore the ambient space dimension n is equal to the number of measurements N.

Figure 3(a) depicts the three receivers and the three different emitter positions. The three receivers are located as follows:

- Receiver 1: (-72.74, -149.86)
- Receiver 2: (-74.05, -71.46)
- Receiver 3: (0, 0)

The four different emitter positions are:

- Emitter 1: (23.32,-49.98)
- Emitter 2: (-17.937,-92.719)
- Emitter 3: (-59.096,-135.554)
- Emitter 4: (-25.364, -72.736)

Figure 4(a) illustrates the comparison of the bias in two different situations: without any bias-correction method (blue bar) and with the proposed bias-correction method (red bar). The standard deviation of the measurement error is 0.05 seconds ($\sigma = 0.05$). Evidently, the proposed method can reduce the localization bias by typically between 70% and 80% for any of the four emitter positions.

The effect of different adjusting the level of noise to 0.1 seconds ($\sigma = 0.1$) and 0.15 seconds ($\sigma = 0.15$) is depicted in figures 4(b) and 4(c) respectively, from which we can conclude that, though the bias is enlarged when the level of noise increases, the proposed bias correction method still performs very well (after reduction, the bias is no more than 0.4 km). The simulation results, demonstrate that the proposed bias correction method is robust to the noise level.

2) Four Receivers Scenario

In this situation, there are four receivers, and we obtain three independent quasi-measurements. We denote these quasi-measurements as t_{12}, t_{23} and t_{34}. The ambient space dimension $n = 2$ is less than the number of quasi-measurements, which corresponds to the $N > n$ situation for the proposed bias-correction, therefore we use the method outlined in Section III to introduce an extra variable which makes $N = n$ again.

Figure 3(b) shows the location of the four receivers and the four emitters positions. The four receivers are positioned as follows:

- Receiver 1: (0, 0)
- Receiver 2: (158.41,-60.21)
- Receiver 3: (158.29,-140.26)
- Receiver 4: (0,-200.24)

There are four different emitter positions:

- Emitter 1: (19.77,-100.12)
- Emitter 2: (79.17,-100.15)
- Emitter 3: (138.58,-100.21)
- Emitter 4: (79.19,-80.08)

Figure 5(a) shows the simulation results in 2-dimensional space with three quasi-measurements. Again, from the figure we can see that the proposed method (the red bars) reduces the bias very effectively (reducing it by up to 75%). Furthermore, by comparing to the simulation results for the three receivers scenario (Figure 4(a)) we can see that introducing an extra sensor, unsurprisingly, improves the accuracy of the localization.

Figures 5(b) and 5(c) illustrate the performance of the proposed bias correction method with different levels of noise ($\sigma = 0.1$ and $\sigma = 0.15$). Similarly to the three receivers scenario, the proposed method is effective in reducing bias even with a high level of noise. Again, comparing to the three receiver scenario (Figure 4(b) and Figure 4(c)) the accuracy of the localization is enhanced by introducing one more sensor.

V. Conclusions

In previous work [5, 6, 7], we proposed a method to determine and thus correct the bias in localization algorithms using range measurements or bearing measurements. The bias-correction method combines Taylor series expansions and Jacobian matrices to express the bias analytically. In this paper we present the bias-correction method in a more generic way. To demonstrate the validity of our bias correction method, we have applied the proposed bias-correction method to the scan-based localization problem in which the original method needs an adjustment, due to the way noise enters the quasi-measurements. Monte Carlo simulation results based on the simulated data provided by DSTO Australia's synthetic integration laboratory (SIL) demonstrated the performance of the proposed bias correction method. Our future work is aimed at improving the performance of the proposed method by using higher order terms of the Taylor series; this may be important in high noise situations.

VI. Acknowledgments

Y. Ji, C. Yu and B.D.O. Anderson are supported by the Australian Research Council under DP-110100538. C. Yu is an ARC Queen Elizabeth II Fellow. B.D.O. Anderson and Y. Ji are also supported by National ICT Australia-NICTA. NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. This material is based on research sponsored by the Air Force Research Laboratory, under agreement number FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government. All the simulated data is provide by Shane Miller who is contracted by the Defence Science and Technology Organization Australia (DSTO).

References

[1] J. L. Melsa and D. L. Cohn. Decision and Estimation Theory. McGraw-Hill Inc., 1978.

[2] J. S. Picard and A. J. Weiss. Localization of Networks Using Various Ranging Bias Models. Wireless Communications and Mobile Computing, 8: pp. 553-562, 2008.

[3] M. Gavish and A.J. Weiss. Performance Analysis of Bearing-only Target Location Algorithms. IEEE Transactions on Aerospace and Electronic Systems, 28(3): pp. 817-827, 1992.

[4] S. P. Drake and K. Dogancay. Some Applications of Tensor Algebra to Estimation Theory. 3rd International Symposium on Wireless Pervasive Computing, ISWPC 2008, pp. 106-110, 2008.

[5] Y. Ji, C. Yu and A. B.O. Anderson. Bias-Correction in Localization Algorithms. IEEE Global Telecommunications Conference, pp. 1-7, 2009.

[6] Y. Ji, C. Yu and A. B.O. Anderson. Localization Bias Correction in n-Dimensional Space. IEEE International Conference on Acoustics Speech and Singal Processing (ICASSP), pp. 578-583, 2010.

[7] Y. Ji, C. Yu and A. B.O. Anderson. Bias-Correction Method in Bearing-only Passive Localization. Processing of European Signal Processing Conference, 2010.

[8] S. M. Kay, Fundamentals of Statistical Signal Processing : Estimation Theory. Englewood Cliffs, N.J. PTR Prentice-Hall, 1993.

[9] R. Shaw. Vector Cross Products in n Dimensions. International Journal of Mathematical Education in Science and Technology, 18(6): pp. 803816, 1987.

[10] H. Hmam. Scan-based Emitter Passive Localization. IEEE Transaction on Aerospace and Electronic Systems, 43(1): pp. 36, 2007.

Localization of Sensor Networks using Bilateration and Four-Bar Linkage Mechanisms

S. Alireza Motevallian, Lu Xia, Brian D.O Anderson

***Abstract*—The paper explores the localization of large-scale sensor networks by splitting them into small sub-networks, localizing these in coordinate bases particular to each subnetwork, and then merging back the results to localize the whole network in a common coordinate basis. Primary focus is on the merging step, and we extend the idea of bilateration in order to merge pairs of sub-networks. This involves the introduction of a novel approach based on the four-bar linkage mechanism. We show that the proposed technique is capable of merging localizations of any pair of sub-networks whose connections are sufficient to ensure the whole network, i.e. subnetworks plus connections, would be in principle localizable using centralized calculations. The main benefit of the algorithm is that it is effectively distributed, with very low computational complexity in comparison with many localization techniques (e.g. MDS-MAP). Another important outcome of this technique is its applicability to a broader class of networks than those treatable with bilateration or trilateration.**

***Index Terms*—Sensor Network Localization, Merging-based localization, Bilateration, Four-bar linkage mechanism**

I. Introduction

In this paper we propose a range-based localization technique where there are a few sensors with known positions (anchors) and there is a set of distance measurements between some pairs of sensors. The network is modeled by a graph called the *grounded* graph (details in Section II-A). Eren et. al. in [5] introduced some conditions (called *global rigidity*) on the grounded graph of the network which ensure its unique localizability. This enables us to study the localizability of the network by using graph theoretical techniques (details in Subsection II-B).

Generally the localization techniques are classified into two classes [2]: centralized (e.g. MDS-MAP [15]) and distributed ([12], [10]). While the Distributed techniques have the benefit of local information exchange and fairly simple calculations over centralized techniques (as discussed in [1] the localization problem is NP-Hard) with the cost of requiring more distance measurements and anchors. In most of the available distributed localization schemes the grounded graph of the network is assumed to be trilaterative, i.e. each node knows its distance to at least three neighbors.

In this work we adopt a merging/splitting strategy in which the whole network is split into small sub-networks each of which yet localizable. These small sub-networks can be efficiently localized with existing techniques in their local coordinate basis and then with the use of an appropriate and efficient merging technique we will be able to find the position of all the nodes in global coordinates. Due to the small size of the sub-networks this technique is totally distributed. A similar idea has been used in [14] to improve the performance of MDS-MAP by computing local maps and then merging them back together.

Trilateration-based techniques are only applicable to dense networks where there are many more known distances than actually required for a network to be uniquely localizable. In [16] it is explained that trilateration techniques fail to localize the boundary nodes while in some sensor network applications like intrusion detection and coverage, these nodes have the most important and valuable information among the sensors. A much sparser and broader class of networks which still can be localized in a distributed manner (i.e. beyond trilateration networks) is the bilateration networks introduced in [8], [7]. In these networks the position of each node is derived from the distances to at least **two** (instead of at least **three** neighbors required for the definition of trilateration networks) neighbors with known positions (details in Section II). Although every trilateration network is a bilateration network, some bilateration networks do not necessarily have unique realizations (no algorithm can find a unique position for every node). Therefore, among these networks we are interested in uniquely localizable ones which we call **bilaterative localizable**.

We combine the idea of splitting/merging with the above mentioned bilateration-based techniques to not only propose a distributed localization technique but also a technique which is applicable to a broader class of networks known as distributedly localizable. The only assumption is that the network is splittable into some localizable sub-networks which may or may not share some nodes, and so that there are enough links between pairs of sub-networks to ensure the localizability of the union. The merging technique proposed here sequentially merges pairs of the sub-networks into larger sub-networks until the whole network results. To accomplish this merging task, we mainly use the idea of bilateration. However, we further enrich it by incorporating the novel idea of four-bar linkage mechanisms from geometry to tackle the possible merging configuration where even the bilateration techniques fail (this configuration is described as unsolvable in [8]).

Section II introduces the background techniques, theories

This research is supported by Australian Research Council Discovery project DP110100538 and by the US Air Force Research Laboratory, under agreement number FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S.Government.

S. Alireza Motevallian and Brian D.O. Anderson are with the Australian National University and National ICT Australia. Lu Xia is with the Australian National University. Emails: {alireza.motevallian, u4488603,brian.anderson}@anu.edu.au

Correspondence to: S. Alireza Motevallian, 115 North Rd., The Australian National University, Canberra, ACT, 0200, Australia. Tel: +61 2 61258639. Fax: +61 2 6125 8660.

and notations we will be using throughout the paper. Section III gives a clear definition of the problem explaining the distinct cases that need to be addressed whereas Section IV provides solutions for them. Section V introduces some further discussions about the algorithm and compares the class of the networks addressed by this algorithm with other classes that have already been solved. We finally give concluding remarks in Section VI.

II. Background

The main focus of the paper is on localization in a 2D ambient space, but wherever it is possible we provide a general definition valid for any d-dimensional ($d \in \{2, 3\}$) ambient space.

A. Network Abstraction

To model a network we first need to introduce the notion of frameworks. The *grounded graph* of the network is a graph $G = (V, E)$ where each vertex $v \in V$ corresponds to a node and there is an edge $(v_1, v_2) \in E$ if the distance between the nodes corresponding to v_1 and v_2 is known. We define $n = |V|$ to simplify the notation. A *configuration* $\boldsymbol{p} = \{p_1, ..., p_n\}$ is a finite collection of n labeled points in $\boldsymbol{R^d}$ where d is the dimension $d \in \{2, 3\}$. A *framework* $F = (G, \boldsymbol{p})$ is a graph $G = (V, E)$ together with a corresponding configuration $\boldsymbol{p} = \{p_1, ..., p_n\}$ in $\boldsymbol{R^d}$. In this framework, $\boldsymbol{p}$ is the mapping $\boldsymbol{p} : V \to \boldsymbol{R^d}$ where each vertex $v_i \in V, i = 1..n$ is assumed to be located at a corresponding point $p_i \in \boldsymbol{p}$. Sometimes this framework is called a realization of the graph G.

In the context of network localization, a network of n sensors therefore can be modeled by the framework $F = (G, \boldsymbol{p})$ where $G = (V, E)$ is the grounded graph of the network and $\boldsymbol{p}$ is a mapping $\boldsymbol{p} : V \to \boldsymbol{R^d}$. We use the terms framework and network interchangeably throughout the paper.

B. Global Rigidity

We say that two frameworks $F_1 = (G, \boldsymbol{p})$ and $F_2 = (G, \boldsymbol{q})$ are *equivalent* (shown by $F_1 \equiv F_2$) if for all pairs i, j where $(i, j) \in E$, $\|p_i - p_j\| = \|q_i - q_j\|$. Two configurations $\boldsymbol{p}$ and $\boldsymbol{q}$ are called congruent, denoted by $\boldsymbol{p} \cong \boldsymbol{q}$ if for all pairs $i, j \in V$, $\|p_i - p_j\| = \|q_i - q_j\|$ [4]. We say that a framework $(G, \boldsymbol{p})$ is globally rigid in $\boldsymbol{R^d}$ if $(G, \boldsymbol{p}) \equiv (G, \boldsymbol{q})$ implies $\boldsymbol{p} \cong \boldsymbol{q}$ for any generic configuration $\boldsymbol{q}$ with $|\boldsymbol{q}| = |\boldsymbol{p}|$. A configuration is generic if the coordinates of the points are algebraically independent from each other. In other terms, a graph G is said to be globally rigid if any set of its equivalent frameworks $(G, \boldsymbol{p})$ can be obtained by the translation, rotation or reflection of the whole framework.

The framework $G(\boldsymbol{p})$ is rigid in $\boldsymbol{R^d}$ if there exists an $\epsilon > 0$ such that for any other configuration $\boldsymbol{q}$ in $\boldsymbol{R^d}$ where $\|\boldsymbol{p} - \boldsymbol{q}\| < \epsilon$, $G(\boldsymbol{p}) \equiv G(\boldsymbol{q})$ implies $G(\boldsymbol{p}) = G(\boldsymbol{q})$. Roughly speaking, a framework $G(\boldsymbol{p})$ is called rigid if by fixing the set of distances corresponding to the edges in G, the formation does not have any flex motions [2].

The following theorem, originally from [6] (Theorem 1), states the necessary and sufficient conditions for a network to be uniquely localizable. By uniquely localizable we mean that it is possible to uniquely identify the position of each node.

Theorem 1. *Suppose that $F = (G, \boldsymbol{p})$ is a framework in $\boldsymbol{R^d}$ modeling a given network of at least $d + 1$ nodes. The framework (and therefore the network it models) F is uniquely localizable if and only if G is globally rigid and there are at least $d + 1$ anchors located in generic positions.*

We say that the graph G is (globally) rigid in $\boldsymbol{R^d}$ if its generic realizations in $\boldsymbol{R^d}$ are (globally) rigid. It is important to mention that either all generic realizations of a graph are (globally) rigid or all are non-(globally) rigid. Therefore, the (global) rigidity property is independent of the configuration of a framework if we restrict the configuration space to only generic ones. Global rigidity has a fully combinatorial characterization in 2D which can be tested with an efficient algorithm [6].

C. Merging localizable sub-networks

We now restrict attention to the case $d = 2$. Merging is the task of introducing a set of links between two localizable networks in order to form a single localizable network. This has been studied in detail in [5], [17]. In terms of global rigidity, this problem is equivalent to merging two globally rigid graphs by adding a set of edges so that the union is a single globally rigid graph. It is possible that the two networks share some nodes. However, as shown in [17], these are special cases of the situation where there is no common node. Therefore, from now on we assume that the sub-networks are node-disjoint unless explicitly specified. Later in Subsection IV-D we address the case where the sub-networks have some common nodes.

In general, a sub-network may have any number of nodes. However, to address the general case, we also assume that each sub-network contains at least 3 vertices. Later in Subsection IV-E, we address the scenario when one of the sub-networks has only 1 or 2 nodes.

The following theorem discusses the necessary (as explained in [17]) and sufficient (as proved in [5]) conditions for such a merging to happen, in terms of global rigidity.

Theorem 2. *Suppose that two disjoint globally rigid graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ with $|V_i| \geq 3$, are connected by a set of links, denoted by L. Then $G = G_1 \cup G_2 \cup G(L)$ is globally rigid if and only if the following two conditions hold:*

1. $|V_i \cap V(L)| \geq 3$, $i = 1, 2$.

2. $|L| \geq 4$.

where $G(L)$ is the graph induced by the edge set L and $V(L)$ is the set of end vertices of the edges in L.

This theorem only provides the conditions to ensure that the merging leads to a globally rigid graph but does not provide a constructive solution to the localization problem. We will address this by providing a distributed algorithm for merging two uniquely localizable sub-networks.

D. Bilateration and Trilateration

Most of the distributed localization techniques proposed so far assume that the underlying network has a trilateration ordering. However, as mentioned earlier, there is a broader class

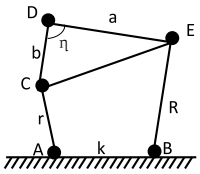


Figure 1. A traditional four-bar mechanism problem setup. A and B are nodes with fixed positions.

of the networks with bilateration ordering (called bilaterative networks), which can be uniquely localized using a distributed algorithm. In addition, the class of trilaterative networks is a proper subset of the class of bilaterative networks. In this section we provide a formal definition for bilateration (trilateration) ordering. A network is said to have bilateration (trilateration) ordering if its grounded graph has such an ordering.

Definition 3. A graph $G = (V, E)$ with $|V| \geq 3$ ($|V| \geq 4$) has a bilateration (trilateration) ordering if its vertices can be ordered as $v_1, .., v_n$ ($n = |V|$) so that the subgraph induced by v_1, v_2, v_3 is a triangle and each $v_i, 3 \leq i \leq n$ ($4 \leq i \leq n$) is connected to *at least* **two** (**three**) vertices v_j, $j < i$ [7].

The **bilateration operation** is the operation of localizing the position of a node up to a set of finite possibilities by considering its distances to at least two neighbors. This can be done by solving the equations obtained from these distance constraints (intersection of multiple circles each one around a neighbor) in addition to the position (global or relative) of those neighbors.

E. Four-bar linkage Mechanism

We will use the idea of four-bar linkage mechanisms in this paper to explore one of the merging scenarios which had been described in [8] as unsolved by bilateration-based localization techniques. The four-bar linkage mechanism has been used in solving a localization problem in [13]. To cope with the space limitations we introduce the mechanism very briefly. Refer to [3] for further details.

Figure 1 shows a typical four-bar linkage. Assume that the vertices A-E are joints and the edges between them are bars with fixed length. If we fix the position of the vertices A and B, and allow the other nodes to move freely in the plane, the position of the node D will follow a curve called a **coupler curve** [3], [13]. Assuming that joint A is the origin and joint B is on the positive side of x-axis, this curve can be explicitly expressed by the degree-six equation (1) where (x, y) is the position of the vertex D [11]:

$$\begin{array}{c} a^2\left((x-k)^2+y^2\right)\left(x^2+y^2+b^2-r^2\right)^2 \\ -2ab\left(\left(x^2+y^2-kx\right)\cos\eta+ky\sin\eta\right) \\ \times\left(x^2+y^2+b^2-r^2\right)\left((x-k)^2+y^2+a^2-R^2\right) \\ +b^2\left(x^2+y^2\right)\left((x-k)^2+y^2+a^2-R^2\right)^2 \\ -4a^2b^2\left(\left(x^2+y^2-kx\right)\sin\eta-ky\cos\eta\right)^2=0 \end{array} \quad (1)$$

III. Problem Statement

Assume that we have a large network which has already been split into small sub-networks with the condition that each of them is uniquely localizable up to at least a local coordinate basis, i.e. the grounded graph of each sub-network is globally rigid. As mentioned before, we assume that each sub-network has at least 3 nodes.

We are interested in a step-by-step merging process in which each sub-network is merged into one of its neighboring sub-networks to form a larger single post-merged localizable sub-network and this merging continues until the whole network results. We will not discuss the conditions required for a network to be splittable in the above manner. However, since any trilaterative network can be treated as special case of merging (as explained in Subsection IV-E if one of the sub-networks contains only 1 node, the merging configuration reduces to a trilateration configuration) and there are merging configurations which are not trilaterative, this class is still strictly larger than the class of trilateration networks (an example is explained in [16]). It is also only required that there are at least 3 anchors somewhere in the network (not necessarily lying in the same sub-network). We also assume that the distance measurements are accurate and there is no noise involving in the estimated distances. One immediate result of these assumptions is that the distances between all pairs of nodes in any localized sub-network are known (we assume that every node in the sub-network knows the location information of the others).

Based on these assumptions, in each step of the algorithm, we take the coordinate basis of a fixed sub-network as the reference coordinate basis and then merge a neighboring sub-network (already localized in its local basis) with this sub-network. This results in a post-merged sub-network (localized in the reference coordinate basis). In further steps, we merge other sub-networks into the already localized (up to the reference basis) sub-networks in a similar manner. The main step of the process (merging two neighboring localizable sub-networks) can then be formally defined as follows:

Assume there are two sub-networks $F_1 = (G_1, \boldsymbol{p_1})$ and $F_2 = (G_2, \boldsymbol{p_2})$ where both have been localized by an arbitrary localization algorithm. F_1 has been localized in the reference basis and F_2 is localized in its local basis. Also assume that there are enough links (and/or common nodes) between F_i satisfying the conditions of Theorem 2. Apply a merging technique to localize the nodes in F_2 into the reference coordinate basis.

If there are three or more vertices in F_2, once the three or more nodes involving in the merging are localized, the other nodes in F_2 can be localized by simply computing the transformation matrix between the reference coordinate basis and the local basis of F_2.

A. Possible configurations

According to Theorem 2, there must be at least **four** distance measurements between F_1 and F_2 in order for a merging algorithm to succeed. It suffices to only consider four links between the F_i ($|L| = 4$) as in the absence of noise, **four** distance measurements are enough to carry out the merging. If there are more than four links, we just pick four of them and proceed with the merging using those links (in Subsection V-A we discuss the use of those extra links to lower the

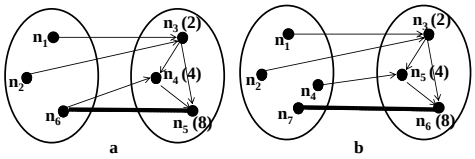


Figure 2. (a) 3-by-3 and (b) 4-by-3 configurations. n_k is the kth node in the bilateration ordering.

computational cost of the algorithm). As it is necessary for each F_i to involve at least **three** vertices, there are 4 possible configurations if the F_i do not share a common vertex:

1) $|V_i \cap V(L)| = 3$;
2) $|V_1 \cap V(L)| = 4$ and $|V_2 \cap V(L)| = 3$;
3) $|V_1 \cap V(L)| = 3$ and $|V_2 \cap V(L)| = 4$;
4) $|V_i \cap V(L)| = 4$.

Allowing the F_i to share some nodes, results in three different scenarios (the solution to these cases will be discussed in Subsection IV-D):

1) $|V_1 \cap V_2| = 1$ and $|L| = 2$ (1 node in common and 2 links joining the F_i);
2) $|V_1 \cap V_2| = 2$ and $|L| = 1$ (2 nodes in common and 1 link joining the F_i);
3) $|V_1 \cap V_2| \geq 3$ and $|L| = 0$ (3 or more nodes in common).

IV. Solution

In this section, we explain how the bilateration technique and four-bar linkage mechanism can be used to merge the F_i. We address the merging for all different configurations case by case, starting from cases where there is no node in common.

A. Three-By-Three

In the first possible configuration, three nodes in each sub-network are joined to each other by the four links in L (Figure 2a). As mentioned before, all the three nodes in F_1 are already localized into the reference basis and acting as anchors. We apply the bilateration operation to localize the nodes in F_2 into the reference coordinates.

Figure 2a shows an example of a three-by-three merging. In this case there is always a node in F_2 which is connected to exactly two nodes in F_1. Applying the bilateration operation (using these links) gives two possible positions for this node. For example in Figure 2a, node n_3 is adjacent to $n_1, n_2 \in V_1$ and therefore using bilateration, we can derive two possible positions for it. Following this figure, the node n_4 is connected to n_3 and n_2 and hence by further applying bilateration, we obtain up to 4 possible positions for it. Following this process, eight possible positions are derived for node n_5 using the connections to n_3 and n_4. However, the distance between n_5 and n_6 is also known, and generic global rigidity of the whole network implies that only one of the eight possibilities is valid. After finding the unique position of n_5, we can backtrack and identify the unique positions of n_4 and n_3 as well, by barring inconsistent position candidates.

B. Four-by-Three & Three-by-Four

Similar to the three-by-three configuration, in the four-by-three case there is a node in F_2 which is adjacent to two nodes in F_1 and therefore can be localized up to a finite number of possible positions using the bilateration operation. Other steps are exactly the same as in the three-by-three case (Figure 2b).

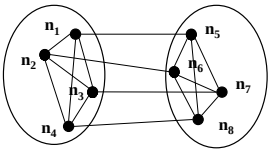


Figure 3. 4-by-4 Configuration. Bilateration fails to extend from left to right.

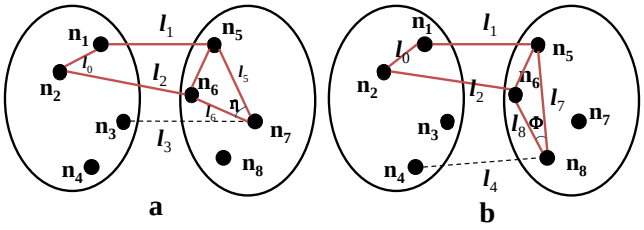


Figure 4. Two possible 4-bar linkages in a 4-by-4 configuration. Each node $n_i, 1 = 7, 8$ lies on the intersection of a coupler curve and a circle.

In the three-by-four case, there is no bilateration ordering which is initiated in F_1 and grows towards F_2. However, with a simple device we can still obtain a bilateration ordering. Instead of starting from F_1, we start from F_2 as if it were the sub-network localized into the reference basis. A similar algorithm as in the four-by-three case will lead to obtaining unique locations of the nodes in F_1 in the coordinate basis of F_2. Since the actual positions of the nodes in F_1 are known a priori (up to the reference coordinates), we can obtain the transformation matrix to convert the local positions of the nodes in F_2 into the reference coordinates.

C. Four-by-Four

In this case four nodes in F_1 and four nodes in F_2 are connected via the four links (Figure 3). Note that in the figure, the edges joining the 4 vertices in F_1 and in F_2 may be virtual edges, i.e. edges whose lengths become known after localizing within the particular F_i. The bilateration technique we have used so far fails here as there is no node in F_2 with more than one neighbor in F_1 and bilateration ordering cannot be extended from F_1 to F_2. In [13], the authors solve a similar problem using the semi-definite programming (SDP) technique. However, this requires a high computational capability which may be beyond the processing capabilities of most of the sensors. Instead, we incorporate the four-bar linkage technique introduced in II-E which can be easily carried out by sensors with low computational capabilities.

1) Four-bar linkage technique: This technique has been previously used in [13] to address some localization problems differing from the subject of this paper. Figure 4 shows how the four-bar linkage mechanism can be used to solve the merging problem. The four links joining the sub-networks of Figure 3 can actually form two separate four-bar linkages as shown in Figure 4. According to [3], the structure in Figure 4a (Figure 4b) leads to a coupler curve for node n_7 (n_8). In both cases, the nodes n_1, n_2 are the fixed bases in the four-bar linkage mechanism.

Assume that n_1 is at the origin and n_2 lies on the positive side of the x-axis. The equation for the position of the node n_7, shown by $K(x, y)$, can be obtained by substituting $k = l_0, r = l_1, R = l_2, a = l_6, b = l_5$ in Equation 1. There is also another coupler curve, called $K'(x, y)$, which can be obtained by mirroring $K(x, y)$ against l_0 ($K'(x, y) = K(x, -y)$).

The position of the node n_7 must also satisfy the circle equation (shown by $C_3(x, y)$) corresponding to link l_3 and the position of n_3:

$$(x - x_3)^2 + (y - y_3)^2 = l_3^2$$

Therefore the position of n_7 must be on the intersection of the two curves $K(x, y)$ and $C_3(x, y)$. The real intersections of the curves can be obtained easily by using any numerical technique (details are removed to cope with space limitation). Although in general, the number of solutions for a system comprising a degree-6 equation and a quadratic equation can be up to 12, surprisingly it has been shown in [3] that the intersection of above equations can have **at most six** possible solutions. Intersecting $K'(x, y)$ and $C_3(x, y)$ also gives at most six other possible positions for n_7. Therefore, we may have a total of maximum **12** possible positions for n_7. We will denote this solution set by $P = p_1, ..., p_s, 2 \leq s \leq 12$.

By a similar procedure, we will obtain a maximum of 12 numerical values for the position of node n_8, denoted by $Q = \{q_1, q_2, ..., q_t\}, 1 \leqslant t \leqslant 12$). Since the distance l_9 between n_7 and n_8 is known, for each $i \in 2, ..., s$ and $j \in 2, ..., t$ we calculate $\|p_i - q_j\|$ and compare it with l_9. If the configuration of the network is generic and the measurements are exact (no noise), then out of these $s \times t$ possibilities for (i, j), there is exactly one pair of positions for n_7, n_8 which is consistent with the distance l_9. Therefore, the distance l_9 will resolve the unique position of nodes n_7 and n_8. These unique positions can help to resolve the unique position of the nodes $n_i, i = 5, 6$ as well.

The apparent downside of this technique is that there may be up to **144** different possible (i, j) pairs which is a large number and the presence of any noise can lead to a wrong decision about the correct positions. However, this is only the worst case theoretical bound and it is important to study the average number of possibilities in real scenarios. We studied this property in random networks, where the positions of the nodes are drawn from a uniform distribution in 2D. To do so, we split the unit square area into two same-sized rectangles by a vertical line. Each side includes 4 randomly located nodes representing $F_i, i = 1, 2$. Then, the coupler and circle equations are derived and solved to obtain the possible positions of n_7 and n_8. This whole process is repeated 20 times to get the average number of possible intersections.

As can be seen in Table I, the average intersection count is **4**. This states that the total number of possibilities for (i, j) is expected to be **16** which is far less than the theoretical worst case of **144** obtained earlier. It is also worth mentioning that in our simulations, we never encountered a case where 12 intersections occurred and the worst case produced was 8 for one node.

Table I
NUMBER OF INTERSECTIONS BETWEEN THE CIRCLE AND THE COUPLER CURVE IN A RANDOM EXPERIMENT OF SAMPLE SIZE 20

Property	mean	min	max	standard deviation
Intersections	4	2	8	1.9

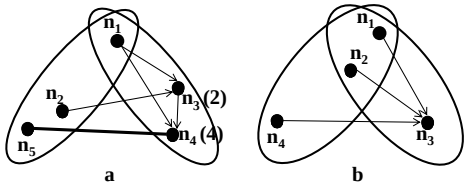


Figure 5. Common vertices: (a) one common vertex; (b) two or more common vertices

D. Sub-networks with Common Nodes

Figure 5a shows an example where 1 node (n_1) is shared between the F_i. Since n_1 is in F_1, n_3 is connected to two vertices from F_1 and therefore we can apply the bilateration operation to find 2 possible positions for it. Again four possible positions are obtained for n_4 using the bilateration operation from the nodes n_1 and n_3. Since the distance between n_4, n_5 is known, it can resolve the unique position of n_4 and with a backtracking the unique position of n_3. If there is more than one vertex in common, the bilateration reduces to a simple trilateration. For example in Figure 5b, n_3 is connected to n_1, n_2, n_4 all from F_1 and therefore can be uniquely localized.

E. Merging a sub-network with 1 or 2 nodes

In this section we provide a solution to the problem of merging when the condition of Theorem 2 is not satisfied and one of the F_i has at most two nodes. Without loss of generality, let us assume that $|F_1| \geq 3$ and $|F_2| \leq 2$. As is shown in [9], a necessary condition for a graph to be globally rigid is that the graph is 3-connected. This implies that $|V_1 \cap V(L)| \geq 3$ (at least three vertices in F_1 must be connected to the vertices in F_2) or otherwise the post-merged graph will not be 3-connected.

F_2 has one vertex: As Figure 6a shows it is enough that the vertex in F_2 is connected to at least 3 vertices in F_1. This is simply the well-known trilateration operation and the intersection of the three circles around the nodes $n_i, i = 1..3$ can uniquely identify the position of node n_4.

F_2 has two vertices: Since, the degree of each vertex $n_i, i = 1, 2$ in F_2 is 1, each n_i must also be connected to at least two vertices in F_1. This simply means that we can apply the bilateration operation starting from F_1 and localize the position of each n_i. Figure 6b shows an example of such a

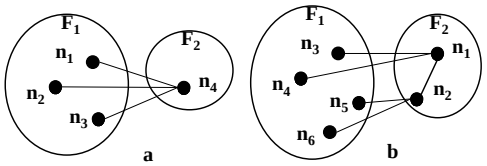


Figure 6. F_2 has 1 or two vertices: (a) $|F_2| = 1$; (b) $|F_2| = 2$

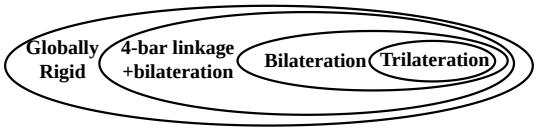


Figure 7. Classification of network topologies localizable by different techniques. Four-bar mechanism enables us to localize networks in which bilateration based algorithms fail.

configuration to explain how the localization algorithm works in this case. First by using the distances from n_3 and n_4 to n_1, we find two possible positions for n_1. Similarly, two positions obtained for n_2 (by using the distances between n_2 and n_5, n_6). Then the actual distance between n_1 and n_2 can be used to resolve their unique positions.

V. Discussions

A. Effect of additional links on computational complexity

In reality, it is often the case that we have more known distance constraints than the minimum required. Therefore, it is relevant to find out the effect of additional links on the computational complexity of the algorithm. However, if there are more links than four (as so far assumed) between the F_i, the chance of having a bilateration ordering extending from F_1 to F_2 obviously increases. Hence, the more links we have, the higher the possibility is that the problem can be solved by bilateration. As bilateration techniques are computationally simpler than the 4-bar linkage algorithm, we conclude that *increasing the number of links between the two sub-networks reduces the complexity of the algorithm with a high probability.*

B. Classification of localizable networks

Figure 7 shows different categories of localizable networks. Trilateration networks have been the target of most of the distributed algorithm as the calculations are fairly simple in these networks. However, these networks are rather dense and might be considered a small subclass of uniquely localizable networks. Bilateration networks on the other hand, form a broader class of networks for which some efficient distributed localization algorithms exist. The downside is that there may be a need for much more memory space than trilateration techniques.

By introducing the four-bar linkage technique, we have addressed a broader class of networks than bilateration which still can be localized by efficient algorithms (Figure 7). Notice that there are still algorithms that can localize all networks with globally rigid grounded graphs but they require **central** calculations.

We should mention that it is still not known that all the networks that can be localized in a distributed manner are localizable by the technique proposed in this paper and this is the subject of further studies.

VI. Conclusion and future work

In this paper we proposed an efficient distributed localization technique based on the idea of splitting large networks into small ones, localizing those small sub-networks and merging them to obtain the wholly localized network. The technique proposed could address all possible merging configurations while keeping the computational complexity very simple. The algorithm adds the idea of four-bar linkage mechanism to bilateration-based techniques in tackling the problem. This led to further broadening the class of the networks that can be distributedly localized and there for outperforming most of the existing distributed techniques which requires the network to be trilaterative.

While we believe that some distributed splitting techniques (e.g. distributed clustering techniques proposed for sensor networks) can perform a splitting suitable for the proposed method in this paper, splitting is still the subject of our future work. We are also studying the effect of the node density of a random network on the success rate of such splitting techniques. In addition, it is still open to investigate the effect of the noise in the accuracy of the technique as if the anchor density is low, it may cause a considerable amount of noise to propagate throughout the network. Finally, it is also presumably possible to extend this algorithm in some way to 3D ambient space, drawing probably on ideas of [17].

References

[1] J. Aspnes, T. Eren, D.K. Goldenberg, A.S. Morse, W. Whiteley, Y.R. Yang, B.D.O. Anderson, and P.N. Belhumeur. A theory of network localization. *IEEE Trans. Mobile Computing*, pages 1663–1678, 2006.

[2] J. Bachrach and C. Taylor. Localization in sensor networks. 2005.

[3] W.Y. Chung. The characteristics of a coupler curve. *Mechanism and machine theory*, 40(10):1099–1106, 2005.

[4] R. Connelly. Generic global rigidity. *Discrete & Computational Geometry*, 33(4):549–563, 2005.

[5] T. Eren, B. Anderson, W. Whiteley, A.S. Morse, and P.N. Belhumeur. Merging globally rigid formations of mobile autonomous agents. In *Proc. 3rd Int. Joint Conf. Autonomous Agents and Multiagent Systems*, volume 3, pages 1260–1261, 2004.

[6] T. Eren, OK Goldenberg, W. Whiteley, Y.R. Yang, A.S. Morse, BDO Anderson, and PN Belhumeur. Rigidity, computation, and randomization in network localization. In *Proc. 23rd Annual Joint Conf. IEEE Computer and Communications Societies, INFOCOM 2004*, volume 4, pages 2673–2684, 2004.

[7] J. Fang, M. Cao, AS Morse, and BDO Anderson. Localization of sensor networks using sweeps. In *Proc. 45th IEEE Conf. Decision and Control*, 2006.

[8] D.K. Goldenberg. *Fine-grained localization in sensor and ad-hoc networks*. PhD thesis, Yale University, 2006.

[9] B. Hendrickson. Conditions for unique graph realizations. *Siam J. Comput.*, 21(1):65–84, 1992.

[10] D. Niculescu and B. Nath. Ad hoc positioning system (aps). In *Global Telecomm. Conf., GLOBECOM'01*, volume 5, pages 2926–2931, 2001.

[11] R.Beyer. *Kinematics Synthesis of Mechanisms*. Chapman and Hall Ltd, 1963.

[12] A. Savvides, H. Park, and M.B. Srivastava. The bits and flops of the n-hop multilateration primitive for node localization problems. In *Proc. 1st ACM int. workshop on Wireless sensor networks and applications*, pages 112–121, 2002.

[13] I. Shames, B. Fidan, B.D.O. Anderson, and H. Hmam. Cooperative self-localization of mobile agents. *IEEE Trans. Aerospace and Electronic Systems*, 47(3):1926–1947, 2011.

[14] Y. Shang and W. Ruml. Improved MDS-based localization. In *INFOCOM 2004. 23rd Joint Conf. IEEE Computer and Communications Societies*, volume 4, pages 2640–2651, 2004.

[15] Y. Shang, W. Ruml, Y. Zhang, and M.P.J. Fromherz. Localization from mere connectivity. In *Proce. 4th ACM int. symp. Mobile ad hoc networking & computing*, pages 201–212, 2003.

[16] Z. Yang, Y. Liu, and X.Y. Li. Beyond trilateration: On the localizability of wireless ad-hoc networks. *IEEE/ACM Trans. Networking (ToN)*, 18(6):1806–1814, 2010.

[17] C. Yu, B. Fidan, and B.D.O. Anderson. Principles to control autonomous formation merging. In *Proc. American Control Conference*, pages 7–pp, 2006.

Noisy localisation on the sphere

Changbin Yu

The Australian National University,
Canberra ACT 2600, Australia
E-mail: Brad.Yu@anu.edu.au

Baoqi Huang*

The Australian National University,
Canberra ACT 2600, Australia
and
National ICT Australia Ltd.,
Tower A, London Circuit, Canberra ACT 2601, Australia
E-mail: Baoqi.Huang@anu.edu.au
*Corresponding author

Hongyi Chee

University of New South Wales,
Sydney, NSW, 2052, Australia
E-mail: Chee.Hongyi@gmail.com

Brian D.O. Anderson

The Australian National University,
Canberra ACT 2600, Australia
and
National ICT Australia Ltd.,
Tower A, London Circuit, Canberra ACT 2601, Australia
and
Research School of Information Sciences and Engineering,
Australian National University,
Building 115, Canberra, ACT 0200, Australia
E-mail: Brian.Anderson@anu.edu.au

Abstract: Localisation is a vital problem in a multitude of research fields, such as navigation, tracking, sensor networks and so on. In previous work, the problem is considered in the plane or in three-dimensional space. This work deals with the problem of distance-based localisation on the surface of the earth when the points lie in a two-dimensional manifold. The challenge lies with finding an appropriate technique to cope with noisy measurements when the conventional formulation for a planar model cannot be used. To this end, we adopt a tool recently applied to the planar model, the Cayley-Menger matrix. Simulation results show that the proposed method is effective and robust to noise. We also quantify the effect of a planar approximation.

Keywords: target tracking; localisation; optimisation; distance geometry; sensor network; planar approximation.

Reference to this paper should be made as follows: Yu, C., Huang, B., Chee, H. and Anderson, B.D.O. (2011) 'Noisy localisation on the sphere', *Int. J. Intelligent Defence Support Systems*, Vol. 4, No. 4, pp.328–350.

Biographical notes: Changbin Yu received his BEng with first class honours in Computer Engineering from Nanyang Technological University, Singapore in 2004 and PhD in Information Engineering from the Australian National University, Canberra, Australia in 2008. He is now a Researcher at the Australian National University and at National ICT Australia. He was a Visiting Researcher at the Technical University of Munich and the National Institute of Informatics, Japan in 2008, Yale University in 2006 and 2007, and Universit Catholique de Louvain, Belgium in 2005. He was a Recipient of an ARC Australian Postdoctoral Fellowship in 2008, the Chinese Government Award for Outstanding Chinese Students Abroad in 2006, and the Australian Governments Endeavour Asia Award in 2005. His current research interests include control of autonomous formations, multi-agent systems, sensor networks and graph theory.

Baoqi Huang received his BEng in Computer Science from Inner Mongolia University, China in 2002 and MS in Software Theory from Peking University, China in 2005. From 2005 to 2008, he worked as a Lecturer in the College of Computer Science, Inner Mongolia University. Currently, he is pursing his PhD in Information Engineering from the Australian National University, Canberra, Australia since 2008. His current research interests include localisation in sensor networks and mobile ad hoc networks.

Hongyi Chee received his BEng from University of New South Wales, Australia in 2007. He contributed to some part of this work as a Summer Scholar at the Australian National University in 2006 to 2007.

Brian D.O. Anderson is a Distinguished Professor at the Australian National University and a Distinguished Researcher at National ICT Australia. He obtained his PhD in Electrical Engineering from Stanford University in 1966. He is a Fellow of the IEEE, IFAC, the Australian Academy of Science, the Australian Academy of Technological Sciences and Engineering, and the Royal Society, and a Foreign Associate of the National Academy of Engineering. His current research interests are in distributed control, including control of UAV formations, sensor networks and econometric modelling.

1 Introduction

Target tracking and navigation have valuable applications in defence, and for any type of target tracking and navigation, target localisation is critical (Danchik, 1988; Healey et al., 1995; Poisel, 2005; Sabour et al., 2008; Tai and Bo, 2009). The goal of localisation is to find the location of a target based on a number of measurements from sensors at known locations. The target can be an unmanned aerial vehicle (UAV), a robot, and even some sort of information that can be sensed. Sensors include vehicles, radar stations, or any other objects at known locations and assisting the localisation of the target. The

measurements may be of different kinds, e.g., distance or time or arrival, time difference of arrival (TDOA), angle of arrival and so on. We consider here the use of distance measurements. Our interest will be localisation on the earth without GPS, with long distances involved. The interesting problems are those where there is noise contaminating the distance measurements. To avoid ambiguities, three or more measurements are needed, and then the question arises of how to allow for the presence of noise in those measurements.

For conventional localisation problems in two- and three-dimensional space, a recent paper (Cao et al., 2006) has shown that an entity formed from the distances between the sensors and the distances between the sensors and the target, termed the Cayley-Menger determinant (CMD), can be used to formulate certain geometric relations among these distances in the noiseless case. This fact can be exploited in the noisy case, so that, as illustrated in Cao et al. (2006), the effect of errors in noisy distance measurements can be reduced, thereby obtaining a better estimate of the target position compared to other approaches to using the noisy distance measurements, see e.g., Niculescu and Nath (2003), Savvides et al. (2003), Savarese et al. (2002), Terwilliger et al. (2004) and Sayed and Tarighat (2005).

However, if observations are made over sufficiently large distances, the surface of the earth cannot be assumed to be flat and the problem accordingly becomes localisation on the sphere (earth); as a result, the CMD approach to localisation with noisy distance measurements cannot be used without some modification. Localisation on the sphere, which has been dealt with in practical applications (Lindsay, 2006; LORAN-C General Information, 1957; Infrasonics Program, 2003; World Wide Lightning Location Network, 2002), often assumes that distance measurements are great circle distances rather than Euclidean distances. For example, in the electronic navigation system LORAN-C used by ships (LORAN-C General Information, 1957), the RF signals transmitted from chains of shore stations at known locations to ships are surface waves propagating on a carrier frequency of 100 KHz out to distances of thousands of kilometres from shore, and closely conform to the earth's curvature; hence, distance measurements are assumed to be great circle distances in a relevant study (Schmidt, 1972). Another example comes from the detection of anomalous HF signals by the Jindalee over-the horizon-radar (OTHR) network. In conjunction with the OTHR a small and widely spaced network of cheap, broadband receivers and spectrum analysers were deployed to localise the anomalous HF signals by cross-correlating these signals to compute TDOA (Lindsay, 2006). As these signals had travelled in a series of hops within the narrow (relative to the earth's radius) wave guide bounded by the ionosphere and the earth's surface and the distance measurements were a significant fraction of the earth's circumference they could be approximated as great circle distances (Newsam, 2006).

In general, localising a target in three-dimensional space requires distance measurements from this target to at least four non-coplanar sensors. In this work, we shall attempt to examine the following problem: "given three sensors and one target on the surface of a sphere, is it possible to localise the target with noisy distance measurements using the CMD method; if so, what kind of performance can we obtain compared to other methods?".

For the purpose of analysis and reducing complexity in this paper, we make the overall assumption that "the earth is a perfect sphere and all sensors and target lie on the

surface of the sphere. We comment briefly near the end of the paper on ways by which one could allow for the ellipsoidal shape of the earth".

The paper is organised as follows. In Section 2, we provide some background information. We state the definition of the CMD in three-dimensional space. In Section 3, we look at how it is possible to formulate a CMD on a sphere with three sensors and one target in the noiseless measurement case as well as in the noisy case. In Section 4, we show how the errors in noisy measurements can be estimated and subsequently reduced by solving an optimisation problem; then we gives the algorithm to locate the coordinates of an unknown point on the surface of the sphere, using noisy or noiseless distance measurements. In Section 5, we investigate the localisation problems under a planar assumption, to identify circumstances where a planar approximation will not produce a large error. The paper ends with concluding remarks and directions for future work in Section 6.

2 Background concepts

The basic problem of distance-based target localisation can be formally defined below: given a set of sensors at known positions, and a set of distance measurements from these sensors to the single unknown target, determine the position of the target. The problem could have many variations: for example, when multiple targets are present, when measurements are noisy, when the sensor positions are noisy (see e.g., Yu, 2007), etc.

2.1 Localisation on the plane using distance measurements

The localisation problem on the plane is simple. In the noiseless case, with two sensors, one can determine the position of a target up to binary ambiguity, and with one extra sensor, uniquely (provided that these sensors are not collinear). When the measurements are noiseless, a conventional multilateration (trilateration in this case) method will solve the problem. One can imagine drawing circles centred at each sensor with radii equal to the associated distance measurement, and determining a common point of intersection. In the absence of sensor collinearity, there is a single such point, being the target position.

In the noisy measurement situation, an additional step has to be performed to compensate the effect of noise. Various approaches have been proposed in Cao et al. (2006), Niculescu and Nath (2003), Savvides et al. (2003), Savarese et al. (2002), Terwilliger et al. (2004) and Sayed and Tarighat (2005). Cao et al. (2006) explored an approach based on using an underlying geometric relationship, expressed using the Cayley-Menger matrix, to formulate a optimisation problem to estimate the noises contained in each of the three sensor to target distance measurements. The proposed method is effective for small noise and when the sensors and/or target are not collinear or close to being collinear. It appeals to an underlying geometric constraint on the true distances. It significantly out-performs methods based on linear calculations, see e.g., Sayed and Tarighat (2005).

In two dimensions, more than three sensors can of course be used. The method of Cao et al. (2006) deals with this. In three dimensions, a minimum of four sensors is required. This is not hard to see, since it is a straightforward generalisation of the two-dimensional case.

2.2 The CMD

The Cayley-Menger matrix of n-points in an m-dimensional space is defined as per (Blumenthal and Gillam, 1943)

$$M = \begin{bmatrix} 0 & d_{0,1}^2 & \cdots & d_{0,n-2}^2 & d_{0,n-1}^2 & 1 \\ d_{1,0}^2 & 0 & \cdots & d_{1,n-2}^2 & d_{1,n-1}^2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ d_{n-2,0}^2 & d_{n-2,1}^2 & \cdots & 0 & d_{n-2,n-1}^2 & 1 \\ d_{n-1,0}^2 & d_{n-1,1}^2 & \cdots & d_{n-1,n-2}^2 & 0 & 1 \\ 1 & 1 & \cdots & 1 & 1 & 0 \end{bmatrix}$$

where $d_{i,j} = d_{i,j}$, $i, j = 0, \ldots, n-1$, $i \neq j$ is the Euclidean distance between the points p_i and p_j. The following is the key theorem which we shall use:

Theorem 2.1: (Blumenthal and Gillam, 1943) Consider an n-tuple of points $p_0,\ldots,p_{n-1}$ in m-dimensional space. If $n \geq m + 2$ then the $(n + 1) \times (n + 1)$ Cayley-Menger matrix $M(p_0,\ldots,p_{n-1})$ has rank $m + 2$.

Given five points in three-dimensional space, this theorem is equivalent to requiring a single relationship among the distances, namely, that the determinant of the matrix M of (1) is zero:

$$\det\left(M\left(p_0, p_1, p_2, p_3, p_4\right)\right) = 0 \tag{1}$$

In the next section, we shall attempt to develop a variant of the concept to use on the sphere. The two-dimensional case has already been discussed in Cao et al. (2006).

3 Formulation of Cayley-Menger constraint on a sphere

3.1 Localisation on a sphere with great circle distances

Given two sensors on a sphere, and noiseless great circle distances to an emitter or target, the target evidently can only be localised with binary ambiguity. On the other hand, measurements from three or more sensors will in general resolve the ambiguity, provided that the sensors are not all located on a common great circle, i.e., they are not coplanar with the centre of the sphere. In the noisy case, similar remarks will apply as for the case of localisation in the plane, and it is clear that one needs a way of handling the noise. Our approach will be first to introduce a Cayley-Menger matrix and determinant appropriate for the sphere, with an analogue of Theorem 2.1 applying to noiseless measurements. Then we will show how to handle the presence of noise.

Consider three sensor nodes 1, 2 and 3, with known positions p_1, p_2, p_3 and a further node 0, the target, with unknown position p_0. All these four points lie on the surface of the sphere. Sensor to target distances on the surface of the sphere in general are given as great circle distances, see Schmidt (1972) and Newsam (2006).

Since a CMD involves the Euclidean distances, it is therefore essential to convert each great circle distance into its corresponding Euclidean distance, or the distance along

the chord formed by the pair of end points. Assuming points on the surface of the sphere are represented by vectors with the sphere centre at the origin, the great circle distance between two points a and b with position vectors $\mathbf{a}$, $\mathbf{b}$ in three-dimensional Euclidean coordinates is obtainable via the following equations (M'Clelland and Preston, 1907):

$$\cos\theta = \frac{<\mathbf{a},\mathbf{b}>}{|\mathbf{a}||\mathbf{b}|} \tag{2}$$

$$\widehat{d_{a,b}} = r\theta \tag{3}$$

Here, <$\mathbf{a}$, $\mathbf{b}$> is the dot product of the position vectors of the two points, r is the radius of the sphere, θ is the angle subtended at the origin, and $\widehat{d_{a,b}}$ is the (accurate) great circle distance between points a and b.

The Euclidean distance d_{ab} can be found by any one of the following equations:

$$d_{ab} = 2r\sin\left(\frac{\theta}{2}\right) = 2r\sin\left(\frac{\widehat{d_{ab}}}{2r}\right) \tag{4}$$

To handle the noisy measurement problem, our first goal is to derive an analogue of a Cayley-Menger determinantal condition, which applies for true distances. However, as noted previously, we would apparently need five points in three-dimensional space to do this. An additional point on the surface of the sphere would resolve the problem, but in practical situations, it would increase the cost of localisation. Adding a ground beacon involves large amounts of infrastructure cost as well as maintenance of the beacons, which does not make sense, if it is just to apply the CMD method. We limit our solution to one with only three sensors and propose the following novel result which is a corollary of Theorem 2.1, simple in retrospect, but perhaps not so obvious until it has been stated:

Corollary 3.1: Let p_0, p_1, p_2, and p_3 be four points on the surface of a sphere of radius r, and suppose that d_{ij} denotes the (exact) Euclidean distance between points p_i and p_j. Then with the definition of the spherical Cayley-Menger matrix (SCM) as

$$SCM = \begin{bmatrix} 0 & d_{01}^2 & d_{02}^2 & d_{03}^2 & r^2 & 1 \\ d_{01}^2 & 0 & d_{12}^2 & d_{13}^2 & r^2 & 1 \\ d_{02}^2 & d_{12}^2 & 0 & d_{23}^2 & r^2 & 1 \\ d_{03}^2 & d_{13}^2 & d_{23}^2 & 0 & r^2 & 1 \\ r^2 & r^2 & r^2 & r^2 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \tag{5}$$

there holds

$$\det(SCM) = 0$$

With the four points p_0, p_1, p_2, p_3, associate a fifth point p_4, which is the centre of the sphere. The Euclidean distance from this fifth point to each of the first four points is r. Therefore, we can form the Cayley-Menger matrix associated with these five points, and it is (5), and because all five points lie in three-dimensional space, the determinant is zero by Theorem 2.1.

Remark 3.1: The problem identified prior to the theorem statement of finding a fifth point is bypassed, by not requiring the fifth point to lie on the surface of the sphere. Choosing it at the sphere centre gives us the relevant distances, including that from the fifth point to the target, whose position though unknown is known to be on the surface of the sphere. We comment in the final section on what might be done when the sphere is replaced by an ellipsoid.

Remark 3.2: One should make the distinction between this definition of *SCM*, as a special case of a three-dimensional CM when four points are co-spherical and one point is the centre of that sphere, with another special CM matrix given in Michelucci and Foufou (2004) for the case when all the five points are co-spherical.

Remark 3.3: It can be easily verified that the determinant of *SCM* becomes the determinant of a CM matrix for a problem in which p_0, p_1, p_2 and p_3 are coplanar when r goes to infinity, i.e., ideas of Cao et al. (2006) are recovered.

3.2 Noisy case: SCM

Let d_{ij} denote the accurate Euclidean distance between nodes i and j with i, $j \in \{0, 1, 2, 3\}$, $i \neq j$. Suppose 0 corresponds to the target, and nodes 1, 2 and 3 to the sensors. If the great circle distance measurements from sensor to target, denoted by $\widehat{d_{0i}}$, are corrupted by noise, we can find the corresponding noisy squared Euclidean distance, denoted by $\overline{\overline{d_{0j}}}^2$, using equation (4). We can postulate the existence of an error variable ϵ_i relating the true values d_{0i}^2 to the noisy values according to

$$\overline{\overline{d_{0i}}}^2 = d_{0i}^2 - \epsilon_i \tag{6}$$

for $i \in \{1, 2, 3\}$. We shall now utilise (6) in conjunction with the geometric constraint condition associated with the SCM.

Substituting $\overline{\overline{d_{0i}}}^2 + \epsilon_i$ in place of d_{0i}^2 in SCM yields a form for the SCM in which noisy measurement values explicitly appear, as the three unknowns ϵ_1, ϵ_2, ϵ_3. Call this form of the matrix *SCM**, to emphasise the dependence on the ϵ_i.

By evaluating the determinant of the matrix *SCM** and setting it to zero, we will then arrive at an equation which will provide a relationship among the errors ϵ_1, ϵ_2, ϵ_3 and it is a relationship which includes the measured data. Whatever the errors are, they must satisfy this relationship. The proof of the theorem is largely parallel to that of Theorem 3 of Cao et al. (2006); however, an important non-coplanarity property has to be argued here.

Theorem 3.4: Let p_0, p_1, p_2 and p_3 four points on the surface of a sphere of radius r, suppose p_1, p_2, p_3 are not coplanar with the centre of the sphere, let d_{0i} and $\overline{d_{0i}}$ denotes the exact and noisy Euclidean distances between points p_0 and p_i, and let ϵ_i denote the associated error between the squares as in (6). Then the errors ϵ_i, $i \in \{1, 2, 3\}$ satisfy a

single algebraic equality which is quadratic though not homogeneous in the ϵ_i's, i.e., for some **A**, **b** and c, there holds

$$\epsilon^T \mathbf{A}\epsilon + \epsilon^T \mathbf{b} + c = 0 \tag{7}$$

where

$$\epsilon = \left[\epsilon_1, \epsilon_2, \epsilon_3\right]^T$$

and where **A**, **b**, c depend on known data and are given in the proof below.

Denote the new matrix as $\overline{SCM}$:

$$\overline{SCM} = \begin{bmatrix} 0 & \overline{d_{01}}^2 + \epsilon_1 & \overline{d_{02}}^2 + \epsilon_2 & \overline{d_{03}}^2 + \epsilon_3 & r^2 & 1 \\ \overline{d_{01}}^2 + \epsilon_1 & 0 & d_{12}^2 & d_{13}^2 & r^2 & 1 \\ \overline{d_{02}}^2 + \epsilon_2 & d_{12}^2 & 0 & d_{23}^2 & r^2 & 1 \\ \overline{d_{03}}^2 + \epsilon_3 & d_{13}^2 & d_{23}^2 & 0 & r^2 & 1 \\ r^2 & r^2 & r^2 & r^2 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Then the determinantal equation yields

$$\det(SCM^*) = \det\left(\overline{SCM}\right) = 0 \tag{8}$$

Partition $\overline{SCM}$ as follows

$$\overline{SCM} = \begin{bmatrix} z_{11} & \mathbf{z_{12}} \\ \mathbf{z_{21}} & \mathbf{Z_{22}} \end{bmatrix}$$

where z_{11} is zero, $\mathbf{z_{12}}$ is a row vector and $\mathbf{z_{21}}$ is a column vector. If $\mathbf{Z_{22}}$ is non-singular, then

$$\det\left(\overline{SCM}\right) = \det\left(\mathbf{Z_{22}}\right)\left[z_{11} - \mathbf{z_{12}}\mathbf{Z_{22}^{-1}}\mathbf{z_{21}}\right] = 0 \tag{9}$$

Observe that $\mathbf{Z_{22}}$ is actually the standard Cayley-Menger matrix associated with the three sensors p_1, p_2, p_3 (lying on the surface of the sphere) and the centre of the sphere. Since these four points are not co-planar by hypothesis, det($\mathbf{Z_{22}}$) is non-zero by a converse of Theorem 2.1, see Michelucci and Foufou (2004). This ensures that $\mathbf{Z_{22}^{-1}}$ exists while z_{11} is zero. Hence, from (9)

$$\mathbf{z_{12}}\mathbf{Z_{22}^{-1}}\mathbf{z_{21}} = 0$$

Given that

$$\mathbf{z_{12}} = \mathbf{Z_{22}^T} = \left[\overline{d_{01}}^2\, \overline{d_{02}}^2\, \overline{d_{03}}^2\, r^2\ 1\right] + \left[\epsilon_1\ \epsilon_2\ \epsilon_3\ 0\ 0\right]$$

we obtain

$$\epsilon^T \mathbf{A}\epsilon + \epsilon^T \mathbf{b} + c = 0$$

where $\epsilon = [\epsilon_1, \epsilon_2, \epsilon_3]^T$, with $\mathbf{A}$ being the top left 3×3 block of $\mathbf{Z}_{\mathbf{22}}^{-1}$,

$$\mathbf{b} = 2\mathbf{A}\left[\overline{d_{01}}^2 \overline{d_{02}}^2 \overline{d_{03}}^2\right]^T$$

and

$$\mathbf{c} = \left[\overline{d_{01}}^2 \overline{d_{02}}^2 \overline{d_{03}}^2 r^2 \ 1\right]\mathbf{Z}_{\mathbf{22}}^{-1} \left[\overline{d_{01}}^2 \overline{d_{02}}^2 \overline{d_{03}}^2 r^2 \ 1\right]^T$$

Remark 3.5: In the event that four or more sensors, or multiple measurements from the same set of three sensors, are available with noisy measurements to the target, one such constraint equation can be found for each selection of three. With N sensors, only $N - 2$ of these constraint equations are independent. One could consider constraint equations using {1, 2, 3}, {1, 2, 4},…,{1, 2, N} for example.

Remark 3.6: The condition that p_1, p_2 and p_3 be coplanar with the centre of the sphere is indeed essential for unique localisation, irrespective of the algorithms used. If the four points were coplanar, there would be two positions for p_0, on each side of the plane, consistent with the distance constraints.

4 Localisation on the sphere

4.1 Determining target location

This subsection explains how to estimate the position of a target from the sensors' positions and the great circle distances from each sensor to the target. Note we have assumed that the sensor positions are accurate and the great circle distance measurements may be noisy.

Consider temporarily the case when the great circle distances are noiseless, i.e., $\widehat{d_{0i}}$ are used. We can write down the following four equations:

$$\frac{\widehat{d_{01}}}{r} = \arccos\left(\frac{\langle \mathbf{p_0} \cdot \mathbf{p_1} \rangle}{|\mathbf{p_0}||\mathbf{p_1}|}\right) \tag{10}$$

$$\frac{\widehat{d_{02}}}{r} = \arccos\left(\frac{\langle \mathbf{p_0} \cdot \mathbf{p_2} \rangle}{|\mathbf{p_0}||\mathbf{p_2}|}\right) \tag{11}$$

$$\frac{\widehat{d_{03}}}{r} = \arccos\left(\frac{\langle \mathbf{p_0} \cdot \mathbf{p_3} \rangle}{|\mathbf{p_0}||\mathbf{p_3}|}\right) \tag{12}$$

$$\mathbf{p_0}^T \mathbf{p_0} = r^2 \tag{13}$$

The set (10) to (13) provides four equations for three unknowns. In the noiseless case, there will exist a unique solution to the equations. Now, suppose that the great circle

distances are noisy. If we simply insert the noisy distances $\widehat{d_{0i}}$ into the equations above, there will in general no longer be any solution to the equations, because they are an overdetermined set.

Let us now indicate an algorithm for obtaining a solution to the equations in the noiseless case, which has the property that if noisy measurements replace noiseless ones in the algorithm, the algorithm can still be executed and it will yield a target position estimate (though not of course one which satisfies (10) through (13) simultaneously, which will be impossible). In the next subsection, we will indicate an improvement to the algorithm for the noisy case.

The algorithm is motivated by what has been suggested for planar localisation with three noisy distance measurements (Sayed and Tarighat, 2005):

1 taking the cosine of both sides of equations (10) and subtracting the transformed (10) from the transformed (12), we can obtain (14)

2 similarly, we can obtain (15) from (11) and (12)

3 these two equations are then combined with (13) and the resulting three equations are solved for the three unknowns.

$$\begin{cases} \cos\left(\dfrac{\widehat{d_{01}}}{r}\right)-\cos\left(\dfrac{\widehat{d_{03}}}{r}\right)=\dfrac{\langle \mathbf{p_0}\cdot\mathbf{p_1}\rangle}{|\mathbf{p_0}||\mathbf{p_1}|}-\dfrac{\langle \mathbf{p_0}\cdot\mathbf{p_3}\rangle}{|\mathbf{p_0}||\mathbf{p_3}|} & (14)\\ \cos\left(\dfrac{\widehat{d_{02}}}{r}\right)-\cos\left(\dfrac{\widehat{d_{03}}}{r}\right)=\dfrac{\langle \mathbf{p_0}\cdot\mathbf{p_2}\rangle}{|\mathbf{p_0}||\mathbf{p_2}|}-\dfrac{\langle \mathbf{p_0}\cdot\mathbf{p_3}\rangle}{|\mathbf{p_0}||\mathbf{p_3}|} & (15)\\ \mathbf{p_0}^T\mathbf{p_0}=r^2 & (16)\end{cases}$$

In the noiseless case, the above method must deliver a correct answer for p_0 due to geometric consistency. This motivates us to utilise the same consistency requirement embedded in the Cayley-Menger determinantal condition to handle the noisy measurements: "one replaces the direct measured noisy great circle distances by a set of estimated great circle distances that have geometric consistency (which is enforced by the use of a CMD constraint) to obtain a target estimate". We can use the optimisation method outlined in the next subsection to obtain estimated great circle distances.

4.2 Optimisation and error reduction

In this subsection, we will see how the errors in the noisy measurements can be estimated, subsequently leading to estimates of the Euclidean distances between the sensors and the target, which are consistent with the geometrical constraint embodied in the CMD being zero. The estimates then allow estimation of the target position.

The analysis is analogous to that in Cao et al. (2006). Let ϵ_i as defined in (6) be the error in the estimated squared distances between the target and sensor *i*. We aim to minimise:

$$J=\epsilon_1^2+\epsilon_2^2+\epsilon_3^2 \tag{17}$$

to the quadratic equality constraint (7). This is actually a reasonably standard problem of numerical analysis. If there happen to be more constraints, on account of having more sensors, the problem is less standard, but nevertheless well posed.

For the single constraint case and using the Lagrangian multiplier method, we obtain the following objective function H:

$$H\left(\epsilon_1,\epsilon_2,\epsilon_3,\lambda_1\right)=\epsilon_1^2+\epsilon_2^2+\epsilon_3^2+\lambda f\left(\epsilon_1,\epsilon_2,\epsilon_3\right) \tag{18}$$

where $f(\epsilon_1, \epsilon_2, \epsilon_3)$ has the quadratic form (7).

By differentiating the objective function H with respect to ϵ_1, ϵ_2, ϵ_3, and λ, and setting the result to zero we can obtain four equations. One of these is (7).

On solving these four equations numerically, we will, often, end up with multiple solutions, with some sometimes being complex numbers. Hence, we need to eliminate any complex solution or non-optimal real stationary point solution, i.e., solutions corresponding to other than the global minimum. The global minimum must be one of the solutions. The solution for the least squares problem is then the set of ϵ_i's which satisfy the condition set out below:

$$\begin{aligned}&\min \epsilon_1^2+\epsilon_2^2+\epsilon_3^2\\&\text{s.t. } f\left(\epsilon_1,\epsilon_2,\epsilon_3\right)=0\end{aligned}$$

As we now have the values for ϵ_1, ϵ_2, ϵ_3, we then obtain the estimated Euclidean distances and subsequently convert them into the estimated great circle distance for each of the sensor to target pairs. The Euclidean distances are consistent with the Cayley-Menger condition, in that if substituted into the determinant, will result in the determinant being zero. This also enforces the geometric consistency of estimated great circle distances, i.e., the four equations (10) to (13) will now have a solution.

In effect, in this subsection we have almost described how to compute a maximum likelihood estimate of the target. We have done this by computing a maximum likelihood estimate of the errors associated with the squares of the Euclidean distances between the sensors and the target which is consistent with the inherent geometrical constraint that links these errors. The entire derivation is very reminiscent of the two-dimensional result of Cao et al. (2006). The argument that Z_{22} is non-singular is peculiar to this problem.

4.3 Computational examples

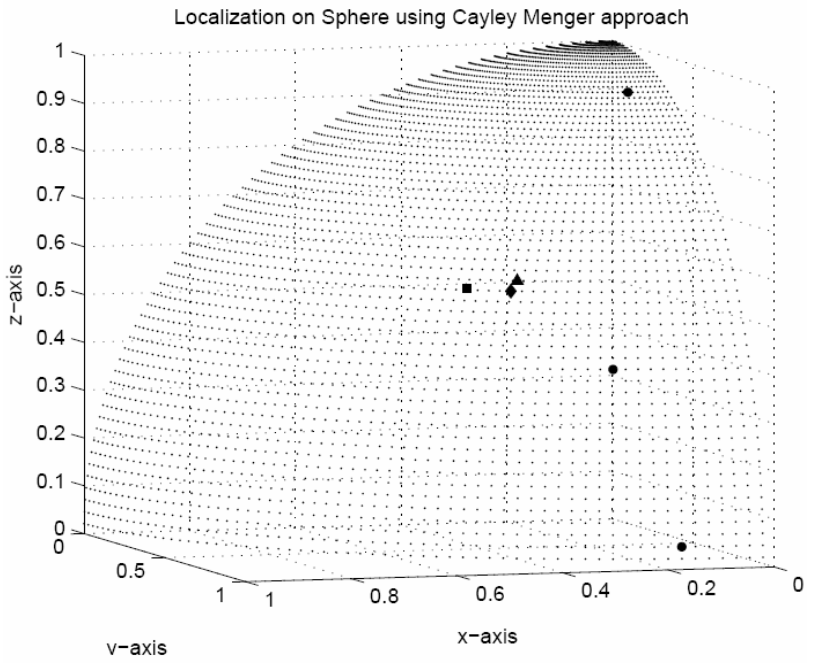
In this subsection, we will give two related computational examples to demonstrate the steps introduced in the previous sections. In the first example, we directly use the noiseless great circle distances from each sensor to the target. While in the second case (noisy), we introduce errors into the great circle distance measurements.

We first describe the setup of both examples, which deal with the same geometric arrangement of sensors and target

- radius of sphere in the two cases is set to unity
- actual target position $\mathbf{p_0}$ = [0.4118, 0.7180, 0.5612], denoted by the diamond in Figure 1

- nodes 1, 2, 3 are designated as the sensors, with known position vectors, denoted by the three-round dots in Figure 1:
  1 sensor position $\mathbf{p_1}$ = [0.0802, 0.3574, 0.9305]
  2 sensor position $\mathbf{p_2}$ = [0.2661, 0.8731, 0.4085]
  3 sensor position $\mathbf{p_3}$ = [0.1698, 0.9844, 0.0458]
  4 the origin serves as a pseudo sensor $\mathbf{p_4}$
  5 for the ten distances that are required in the three-dimensional Cayley-Menger matrix, only the three great circle distance measurements between target and sensors, namely $\left[\widehat{d_{01}}, \widehat{d_{02}}, \widehat{d_{03}}\right]$ will change, as a result of varying the target position $\mathbf{p_0}$.

Figure 1 Noisy case: round dots denote sensors, diamond denotes actual target, triangle denotes optimised target estimate based on using SCM-optimised distance estimates, square denotes a target estimate based on using noisy distances directly



In the first example, we shall show the performance of the CMD method in a noiseless situation by using the matrix *SCM* of Section 3. The noiseless great circle distances measurements between the actual target and sensor positions are:

$$\widehat{d_{01}} = 0.6236,\ \widehat{d_{02}} = 0.2627,\ \widehat{d_{03}} = 0.6395$$

By converting the values of $\widehat{d_{0i}}$ into their corresponding Euclidean distances d_{0i} using (4), we can then substitute the corresponding Euclidean distances into the matrix *SCM*. Evaluating the determinant of *SCM* results in a value of 0 (as expected).

This indicates that the Euclidean distances are consistent with the set of points (p_0, p_1, p_2, p_3, p_4) in three-dimensional space and hence there is no correction needed. As independent verification of this, we note that solutions for the errors obtained from MATLAB simulation using the algorithm in Subsection 4.2 are as follows:

$$\epsilon_1^* = 0.000,\ \epsilon_2^* = 0.000,\ \epsilon_3^* = 0.000$$

Using these accurate great circle distances and following the method in Subsection 4.1, we obtain the estimated target position $\mathbf{p_0'} = [0.4118, 0.7180, 0.5612]$, which is the same as the true target position as expected.

Let us now consider the case where the measurements are noisy. All nodes and parameters in this case remain the same as those in the first case apart from the three great circle distance measurements from sensors to the target. Suppose the three great circle distance measurements are corrupted by noises within [–4%, 4%] of the actual distances as following

$$\widetilde{d_{01}} = 0.5986$$

$$\widetilde{d_{02}} = 0.2575$$

$$\widetilde{d_{03}} = 0.6587$$

By converting each great circle distance $\widetilde{d_{0i}}$ into its corresponding Euclidean distance $\overline{d_{0i}}$, we can then substitute the values of $\overline{d_{0i}}$ into the matrix *SCM**. By evaluating the determinant of the matrix *SCM**, we obtain one quadratic equality constraint, defined by (7). Subsequently, we followed the procedures as mentioned in Subsection 4.2 and obtained the following solution to the constrained least squares problem: $\epsilon_1^+ = -1.235\times10^{-3},\ \epsilon_2^+ = 7.886\times10^{-3},\ \epsilon_3^+ = -4.216\times10^{-3}$.

Following the method in Subsection 4.1, we can solve for the optimised target estimate and we obtain the position [0.3973, 0.7099, 0.5816].

For a direct comparison with a non-optimised estimate, the noisy great circle distance measurements $\widetilde{d_{01}}$, $\widetilde{d_{02}}$ and $\widetilde{d_{03}}$ are used directly in (10) to (13) and we obtain a target estimate at position [0.4804, 0.6713, 0.5643] whose error is clearly substantial.

As depicted in Figure 1, for this example which is reasonably generic, the CMD method of estimating errors results in a better estimation of the unknown target location on the surface of the sphere as compared to just using the noisy measurements for localisation without utilising the geometric constraints.

5 Localisation under a planar assumption

As we have mentioned in Section 2, if distances between the nodes including sensors and a target are small, the surface of the earth involving these nodes is almost flat so that the two-dimensional CMD method will work approximatively under a planar assumption. However, besides errors from noisy measurements, new errors are induced due to the

planar approximation arising from the planar assumption. With the scale of the distances rising, the errors from the planar approximation will increase and eventually become unacceptable. Therein, the planar assumption does not hold and we can apply the spherical CMD method instead of the two-dimensional CMD method. In this section, we shall investigate the circumstances where the two-dimensional CMD method can be employed under a planar assumption.

5.1 Problem model

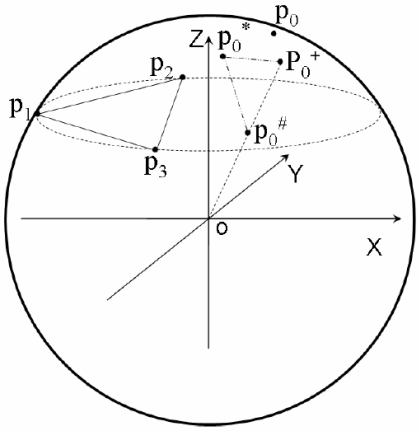
Figure 2 illustrates the target localisation problem solved by both the two-dimensional CMD method and the spherical CMD method:

- point p_0 represents the target to be localised and points p_1, p_2, p_3 represent the three sensors with known positions, all which lie on the surface of a sphere corresponding to the earth
- points p_0^* and $p_0^\#$ denote estimated positions of the target p_0 by using the spherical CMD method and the two-dimensional Cayley-Menger under a planar assumption respectively
- point p_0^* lies on the surface of the sphere and point $p_0^\#$ lies in the plane, named $\mathcal{A}$, which is determined by points p_1, p_2 and p_3
- points lying both in $\mathcal{A}$ and inside the sphere form a disk, named $\mathcal{C}$. Note that $p_0^\#$ may be inside, on, or outside the sphere, and thus inside or outside $\mathcal{C}$
- point p_0^+ is the intersection between the surface of the sphere and the ray starting from the centre of the sphere, denoted O, and going through point $p_0^\#$, and it is the projection of point $p_0^\#$ onto the surface of the sphere.

A Cartesian coordinate system is established in Figure 2 with the centre O of the sphere as the origin, Z-axis perpendicular to $\mathcal{A}$ and arbitrary orthogonal X- and Y-axes in a plane parallel to $\mathcal{A}$. Designate $\mathbf{p}$ as the position vector of point p in the coordinate system. We further define the following:

- $d = \max\{d_{01},\ d_{02},\ d_{03},\ d_{12},\ d_{13},\ d_{23}\}$
- $d_{A_{\min}} = \min\{d_{12},\ d_{13},\ d_{23}\}$ and $d_{A_{\max}} = \max\{d_{12},\ d_{13},\ d_{23}\}$
- let $\tilde{\sigma}$ be the great circle distance between points p_0^* and p_0^+ and σ be the Euclidean distance between them, i.e., $\sigma = \left|\mathbf{p_0^*} - \mathbf{p_0^+}\right|$
- $l = \left|\mathbf{p_0^\#}\right|$
- r_p is the radius of the disk $\mathcal{C}$
- h is the distance from the origin O to $\mathcal{A}$.

Figure 2 A localisation scenario



Notes: The solid circle denotes a great circle of the sphere. The dashed ellipse denotes the boundary of the disk $\mathcal{C}$. $\tilde{\sigma}$ is the great circle distance between points p_0^* and p_0^+ and σ is the Euclidean distance between p_0^* and p_0^+.

Recall that the spherical localisation procedure cannot be applied if the three sensors are on a common great circle. Therefore, we impose the practical restriction that they cannot lie in a plane that is closer than αr to the centre of the sphere, i.e.,

$$\frac{h}{r} > \alpha \tag{19}$$

and acceptable values for α will be indicated below.

As depicted in Figure 3, with the least internal angle $\angle p_2 p_1 p_3$ of the triangle formed by the three sensors goes to 0, $\angle p_2 O_p p_3$ goes to 0 at the same time and points p_2 and p_3 tend to overlap; consequently, the three sensors tend to be collinear, with the result that the 2-dimensional localisation procedure does not work. Since the angle $\angle p_2 O_p p_3$ approximately equals to $\frac{d_{A_{\min}}}{r_p}$, we impose another practical restriction that $\angle p_2 O_p p_3$ is larger than a small constant β, i.e.,

$$\frac{d_{A_{\min}}}{r_p} > \beta \tag{20}$$

and again, acceptable values for β will be indicated below.

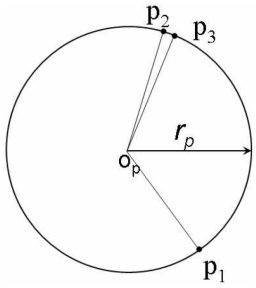
As $\frac{d_{A_{\min}}}{r_p}$ approaches 0, the three sensors tend to be collinear and therefore also to be coplanar with the origin O, which means $\frac{h}{r}$ falls to 0 as well. Thus, the values of α and β are not independent. In addition, when $d_{A_{\min}}$ and r_p simultaneously approach 0, it is possible for the least internal angle to be still larger than β while h approaches r (certainly larger than αr), but the three sensors gradually concentrate at one point, which results in

both collinearity and coplanarity of the three sensors and the centre of the sphere. To avoid the exceptional case, we restrict the minimal inter-sensor distance $d_{A_{\min}}$ to be larger than a small positive constant ρ, i.e.,

$$d_{A_{\min}} > \rho \tag{21}$$

which can be easily fulfilled. For instance, $\rho = 0.001$ km is reasonable in a real system but sufficient for the constraint.

Figure 3 Nearly collinear sensors



Note: The circle corresponds to $\mathcal{C}$ with O_p as its centre and radius r_p.

Essentially, both α and β are determined by the geometric layout of three sensors involved in a localisation problem, and together with ρ describe how close the localisation problem is to the unacceptable situations, i.e., the three sensors being collinear and the three sensors being coplanar with the centre of the sphere. Both of them are suggested by the simulation evidence of Section 5.3 as being separately necessary lower bounds.

5.2 The error from a planar approximation

Measuring the difference between localisation results for one target with and without a planar assumption, namely the Euclidean distance between the two position estimates $\left|\mathbf{p}_\mathbf{0}^* - \mathbf{p}_\mathbf{0}^\#\right|$, defines a metric for the error from the planar approximation. In Figure 2, we notice that the great circle distance between p_0^* and p_0^+, i.e., $\tilde{\sigma}$, and the Euclidean distance between $p_0^\#$ and p_0^+, i.e., $\left|\mathbf{p}_\mathbf{0}^\# - \mathbf{p}_\mathbf{0}^*\right|$, are like two components of $\left|\mathbf{p}_\mathbf{0}^* - \mathbf{p}_\mathbf{0}^\#\right|$ along two different directions: 'angular' and 'radial'. Moreover, the great circle distance $\tilde{\sigma}$ reflects the aspect of the error that directly affects prospective usage of the estimated position under the planar approximation. The Euclidean distance $\left|\mathbf{p}_\mathbf{0}^\# - \mathbf{p}_\mathbf{0}^*\right|$ reflects the aspect of the error from the non-planar characteristic of the spherical surface, and it can easily be compensated for by a projection step following the two-dimensional localisation. Thus, the angular error is of great importance.

Because the Euclidean distance σ between p_0^* and p_0^+ approaches the corresponding great circle distance $\tilde{\sigma}$ when $\tilde{\sigma}$ is small, we use σ as an approximation to $\tilde{\sigma}$. When $p_0^\#$ lies inside the sphere as illustrated in Figure 2, $\left|\mathbf{p}_\mathbf{0}^\# - \mathbf{p}_\mathbf{0}^+\right| = r - 1$. But there are two other

cases according to the position of $p_0^{\#}$. One is that $p_0^{\#}$ lies on the sphere, where $r = 1$ and $\left|\mathbf{p_0^{\#}} - \mathbf{p_0^{+}}\right| = 0$; the other is that $p_0^{\#}$ lies outside the sphere, where $\left|\mathbf{p_0^{\#}} - \mathbf{p_0^{+}}\right| = l - r$. Thus, we conclude that $\left|\mathbf{p_0^{\#}} - \mathbf{p_0^{+}}\right| = |r - 1|$. As such, we can use σ and $|r - 1|$ to denote the errors from the angular and the radial directions respectively. Moreover, because d indicates the scale of distances involved in the localisation problem, it is appropriate for it to be used to normalise the errors, i.e., dividing the absolute distances by it. Therefore, instead of using the metric $\left|\mathbf{p_0^{*}} - \mathbf{p_0^{\#}}\right|$ for the error introduced by the planar approximation, we define two *sub-errors*, i.e., the angular error and the radial error, as follows

$$e_a = \frac{\sigma}{d} \tag{22}$$

$$e_r = \frac{|r - l|}{d} \tag{23}$$

As in Huang et al. (2008), the notation $O(\cdot)$ is used to describe orders of magnitude of some quantities. Suppose in particular that f is a function of variable x. For some interval $\mathcal{I}$ of $\mathcal{R}$, typically including 0 or ∞, $f = O(x)$ means for some constant k, there holds $|f| \le k|x|$ for all x in $\mathcal{I}$. We can also extend the definition to treat powers of x. In Huang et al. (2008), we obtain orders of magnitude of sub-errors as follows:

$$e_a = O\left(\left(\frac{d}{r}\right)^3\right) \tag{24}$$

$$e_r = O\left(\frac{d}{r}\right) \tag{25}$$

which show that e_r is roughly proportional to d and the power in the order of e_a is cubic rather than linear. Since there is always a possibility of compensating for e_r through a projection operation, we are more concerned with e_a.

5.3 Simulations

The above conclusions express errors in terms of orders of magnitude of certain quantities. In this section, we provide simulation evidence for the analytical results which also allows us to make more precise statements about the levels of error. These simulations treat a large number of different localisation problems involving three sensors and one target, with MATLAB used to determine localisation solutions. In all instances, the same problem is solved by using both the two-dimensional and spherical CMD methods. Further details of the simulations are as follows:

- r is assigned to be 6,371.3 km, which is the average radius of the earth
- the position of every node, namely three sensors and one target in each instance, is random, being obtained by generating three spherical coordinates, a constant radial

distance r, a random zenith angle and a random azimuthal angle (if the values of parameters d, α, β and ρ are required to fulfil certain constraints, the positions are regenerated until the constraints are fulfilled)

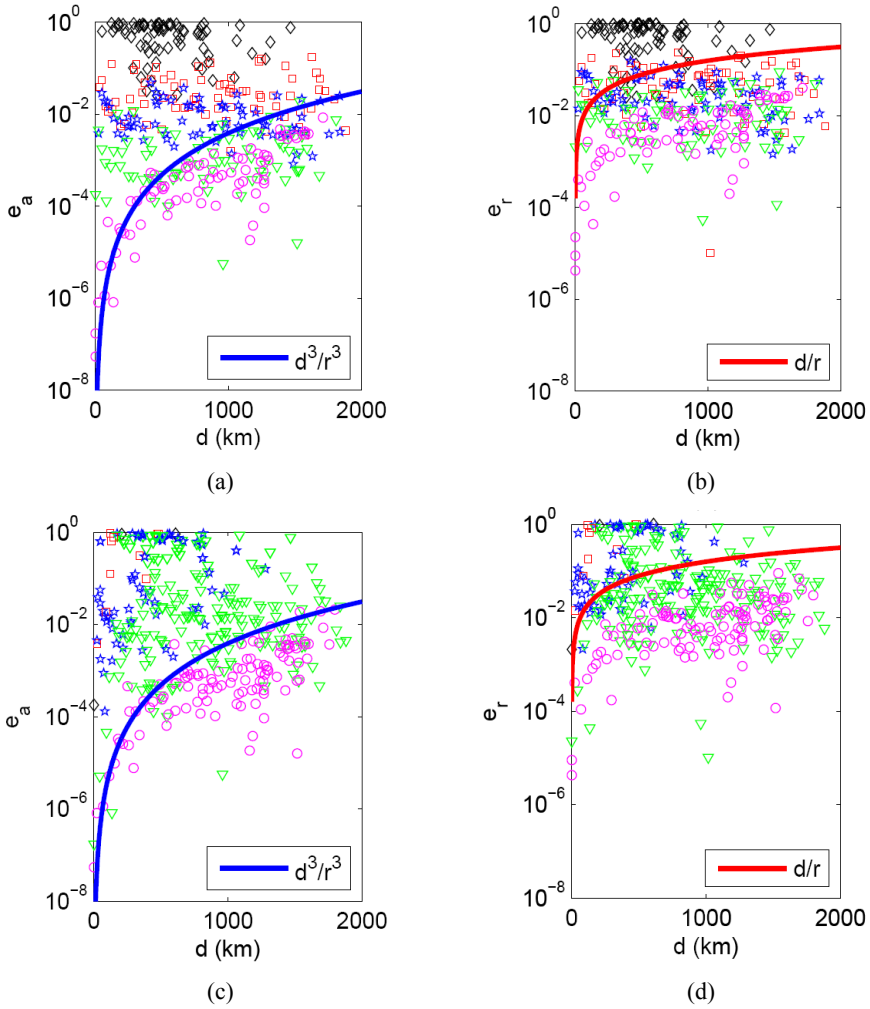
- errors in distance measurements are independent and with different noise levels, see below.

At first, we conduct simulations with zero noise to probe the effects of α and β on the error and plot the sub-errors e_a and e_r in Figure 4. As can be seen from Figure 4(a) and 4(b), it is evident that different scales of $\frac{h}{r}$ result in different levels of the sub-errors except for $\frac{h}{r} > 0.995$: when $\frac{h}{r} \le 0.1$, both sub-errors are nearly as large as 100% no matter what d is; but when $\frac{h}{r} > 0.995$, they are small and increase with d. The Figure 4(c) and 4(d) show that the instances with large sub-errors (corresponding to the marks on top of the figures) generally have small values of $\frac{d_{A_{\min}}}{r_p}$. When $\frac{d_{A_{\min}}}{r_p} > 0.1$, both sub-errors are relatively small and increase with d.

The above simulations show that for the same value of d, small values of $\frac{h}{r}$ and $\frac{d_{A_{\min}}}{r_p}$ normally lead to large sub-errors. The principal reason is that a small value of $\frac{d_{A_{\min}}}{r_p}$ or $\frac{h}{r}$ implies that the three sensors tend to be either collinear or coplanar with the pseudo-sensor, which probably causes a large error in the estimated position, and thus the difference between the estimated positions by using the two-dimensional and spherical CMD methods probably rises as well. Therefore, based on the simulations, we suggest that the parameters α and β should be no less than 0.995 and 0.1 respectively to derive acceptable results when applying a planar approximation.

The case of $\alpha = 0.995$ indicates $r_p < 636$ km, and because $d_{A_{\max}}$ denotes the maximal inter-sensor distance which is less than the diameter of $\mathcal{C}$, we have $d_{A_{\max}} < 1{,}272$ km. On the other hand, the case of $\beta = 0.1$ indicates that the least angle formed by point O_p and any two of the three points p_1, p_2, p_3 is larger than 2.86°. In addition, when $\frac{h}{r} > 0.995$ and $\frac{d_{A_{\min}}}{r_p} > 0.1$, both sub-errors appear to increase with d and their overall growing trends are consistent with the curves corresponding to $\frac{d^3}{r^3}$ and $\frac{d}{r}$ respectively, which also verifies equations (24) and (25).

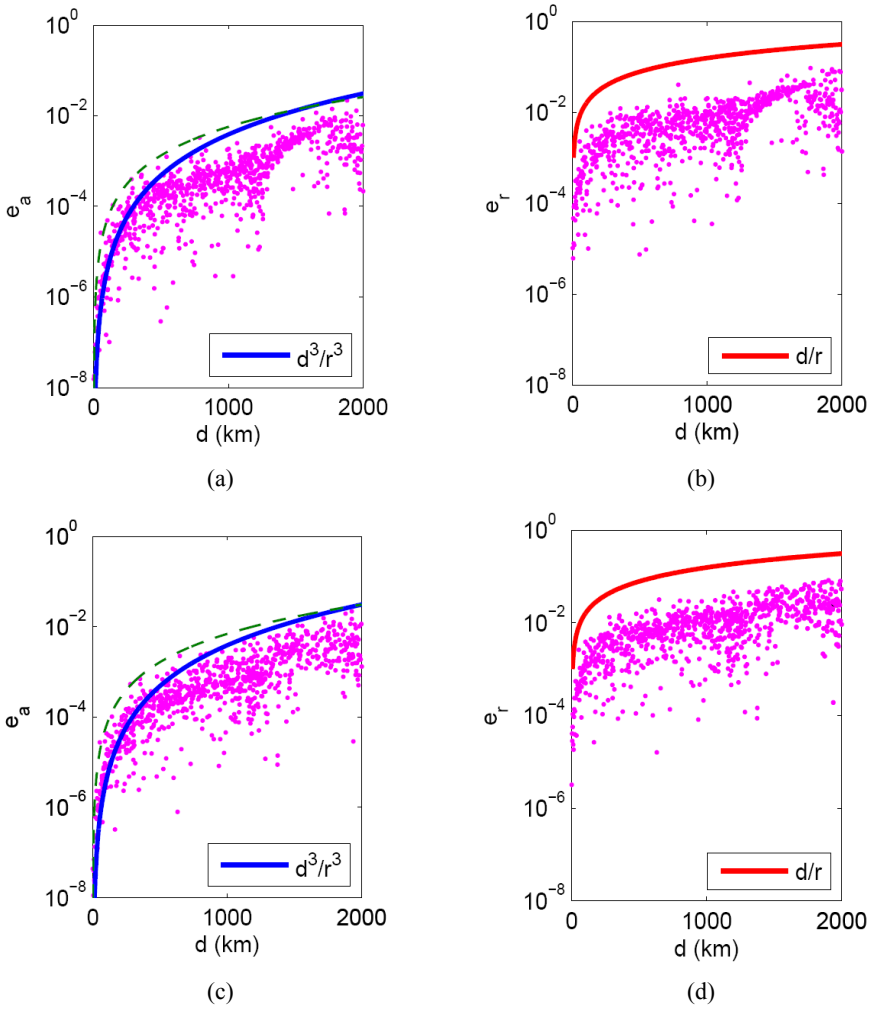
Figure 4 Simulations with zero noise, in (a) and (b), sub-errors e_a and e_r with different scales of $\frac{h}{r}$ are displayed in different shapes: diamond, $0 < \frac{h}{r} \le 0.1$; square, $0.1 < \frac{h}{r} \le 0.8$; pentagon, $0.8 < \frac{h}{r} \le 0.9$; triangle, $0.9 < \frac{h}{r} \le 0.995$; circle, $0.995 < \frac{h}{r} < 1$, in (c) and (d), sub-errors e_a and e with different scales of $\frac{d_{A_{\min}}}{r_p}$ are displayed in different shapes: diamond, $0 < \frac{d_{A_{\min}}}{r_p} \le 0.0001$; square, $0.0001 < \frac{d_{A_{\min}}}{r_p} \le 0.001$; pentagon, $0.001 < \frac{d_{A_{\min}}}{r_p} \le 0.01$; triangle, $0.01 < \frac{d_{A_{\min}}}{r_p} \le 0.1$; circle, $0.1 < \frac{d_{A_{\min}}}{r_p} < \sqrt{3}$ (see online version for colours)



Note: The vertical axes are in logarithm scale.

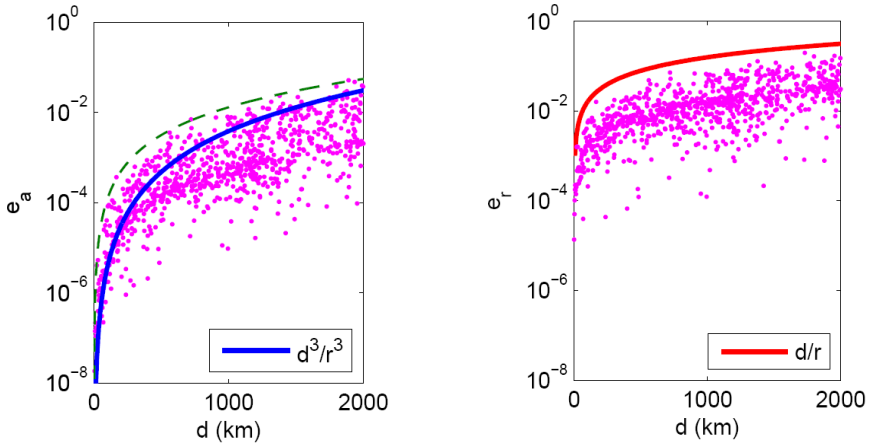
Secondly, we simulate the same localisation problems for zero noise and two different noise levels: 10% and 30% (x% means that the percentage of the distance measurement error is uniformly distributed within $-x$% and x%) and the parameters α, β, and ρ are assigned to be 0.995, 0.1 and 0.001 km respectively. The resultant sub-errors e_a and e_r are plotted in Figure 5. The overall growing trends of both sub-errors are still near the solid curves corresponding to $\frac{d^3}{r^3}$ and $\frac{d}{r}$. Moreover, both sub-errors increase slightly with noise rising from zero to 10% but increase apparently with noise rising from 10% to 30%.

Figure 5 Simulation for different levels of noise when $\alpha = 0.995$, $\beta = 0.1$ and $\rho = 0.001$ km, (a), (c) and (e) show the sub-error e_a (b), (d) and (f) show the sub-error e_r (see online version for colours)



Notes: Distance measurements in (a) and (b) are noiseless; distance measurements in (c) and (d) are with 10% noise; distance measurements in (e) and (f) are with 30% noise. Each dot represents a sub-error in one localisation instance. The vertical axes use a logarithm scale.

Figure 5 Simulation for different levels of noise when $\alpha = 0.995$, $\beta = 0.1$ and $\rho = 0.001$ km, (a), (c) and (e) show the sub-error e_a (b), (d) and (f) show the sub-error e_r (continued) (see online version for colours)



(e) (f)

Notes: Distance measurements in (a) and (b) are noiseless; distance measurements in (c) and (d) are with 10% noise; distance measurements in (e) and (f) are with 30% noise. Each dot represents a sub-error in one localisation instance. The vertical axes use a logarithm scale.

By the method of trial and error, we can derive three approximate upper bounds for the three noise levels on e_a : $\frac{d^3}{5r^3}+\frac{d^2}{5r^2}$, $\frac{d^3}{5r^3}+\frac{d^2}{4r^2}$ and $\frac{d^3}{5r^3}+\frac{d^2}{2r^2}$, corresponding to the dashed curves in Figure 4. Table 1 lists some values of these upper bounds of e_a. Since we are more concerned with e_a, the data in Table 1 give us great confidence in applying the planar approximation due to their small magnitude.

Table 1 Upper bounds on e_a with zero, 10% and 30% noise

d(km)	*10*	*600*	*1,000*	*2,000*
Zero	4.93*e* − 5%	0.19%	0.57%	2.59%
10%	6.17*e* − 5%	0.24%	0.69%	3.08%
30%	1.23*e* − 4%	0.46%	1.31%	5.55%

Whether a planar approximation is acceptable is decided by multiple factors, including accuracy requirements on estimated positions, noise levels of distance measurements and the characteristics of the localisation problems described by parameters, such as d, α, β, and ρ. Given certain parameters, we can predict the upper bounds on both sub-errors e_a and e_r, compare the upper bounds with the accuracy requirements on position estimates and uncertainties in distance measurements, and then decide whether the planar approximation is acceptable. For example, provided that in a sensor network the parameters α, β, and ρ are the same as we have assigned in simulations corresponding to Figure 5 and the noise level is 10%, e_a will be trivial (less than 0.69%) as long as d is less

than 1,000 km. Otherwise, if sensors are equipped with exact distance measuring devices and accordingly high accuracies on positions are required, or an extremely small value of one parameter is admitted, we should be more cautious in accepting the planar approximation.

6 Concluding remarks

In this paper, we have dealt with the problem of localisation on earth using distance measurements, when the planar assumption becomes invalid. We use the geometrical constraints for compensation of the effect of noisy measurements, by formulation of a SCM. Although three-dimensional ideas are being used (for which normally four sensors would be expected), localisation can be achieved with but three sensors, as for the case of planar localisation. This simple yet effective idea is verified using simulation examples. The *numerical range of validity* of a planar assumption for practical problems of localisation on the earth's surface is also studied. Clearly for small enough distances, a planar approximation will be satisfactory.

A number of issues remain to be addressed. The SCM is expressed using Euclidean distances for simplicity and following convention. An expression using great circle distances (and trigonometric functions) is an easy extension. A different variation is obtained by adding altitude measures into the formulation; the problem can be easily extended to the case when the sensors and the target are at different heights, though some a priori estimate of target height would be required, to be able to record a Euclidean distance from the centre of the earth. Note that ellipsoidal models are in fact available for certain large areas of the earth, e.g., of the size of Australia; one could imagine an iterative localisation scheme, in which a height value was assumed based on the current position estimate, and then a new position estimate would be obtained.

A projection method might be used to localise points on sphere. Sensor to target distance measurements could be projected onto the plane formed by the three sensors, and then a planar model can be used. However, the complexity and the effectiveness of this approach are yet to be determined.

Acknowledgements

This work is in part supported by National ICT Australia-NICTA, which is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. C. Yu is supported by the Australian Research Council through an Australian Postdoctoral Fellowship under DP-0877562. The authors are grateful to Ming Cao for the assistance and discussions he kindly provided.

References

Blumenthal, L.M. and Gillam, B.E. (1943) 'Distribution of points in n-space', *The American Mathematical Monthly*, Vol. 50, No. 3, pp.181–185.

Cao, M., Anderson, B.D.O. and Morse, A.S. (2006) 'Sensor network localization with imprecise distances', *System and Control Letters*, Vol. 55, pp.887–893.

Danchik, R.J. (1988) 'Navy navigation satellite system status', in *Position Location and Navigation Symposium*, Record, Navigation into the 21st Century, IEEE PLANS '88., IEEE, pp.21–24, November–2 December.

Healey, A.J., Pascoal, A.M. and Pereira, F.L. (1995) 'Autonomous underwater vehicles for survey operations: theory and practice', in *American Control Conference, Proceedings of the*, June, Vol. 5, pp.2943–2949.

Huang, B., Yu, C. and Anderson, B.D.O. (2008) 'Noisy localization on the sphere: planar approximation', in *Proc. of the 4th International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, Sydney, Australia, pp.49–54.

Infrasonics Program, Earth System Research Laboratory, National Oceanic and Atmospheric Administration (2003) Available at http://www.esrl.noaa.gov/psd/programs/infrasound.

Lindsay, L.M. (2006) 'Geolocation of hf signals using time difference of arrivals methods', in *Workshop on the Application of Radio Science (WARS2006)*, Leura, NSW, Australia.

LORAN-C General Information, Navigation Center, US Coast Guard (1957) Available at http://www.navcen.uscg.gov/?pageName=loranMain.

M'Clelland, W.J. and Preston, T. (1907) *A Treatise on Spherical Trigonometry with Applications to Spherical Geometry and Numerous Examples*, Macmillian and Co. Limited.

Michelucci, D. and Foufou, S. (2004) 'Using Cayley-Menger determinants for geometric constraint solving', in *Proc. of the 9th ACM Symposium on Solid Modelling and Applications*, pp.285–290.

Newsam, G.N. (2006) 'Locating events on the surface of the sphere from time-of-arrival data', in *Proc. Defence Applications of Signal Processing*, Fraser Island, Queensland, Australia.

Niculescu, D. and Nath, B. (2003) 'Localized positioning in ad hoc networks', *Elseviers Journal of Ad Hoc Networks, Special Issue on Sensor Network Protocols and Applications*, Vol. 1, pp.247–259.

Poisel (2005) *Electronic Warfare Target Location Methods*, Artech House Publishers.

Sabour, N.A., Faheem, H.M. and Khalifa, M.E. (2008) 'Multi-agent based framework for target tracking using a real time vision system', in *Computer Engineering and Systems*, *ICCES 2008*, *International Conference on*, November, pp.355–363.

Savarese, C., Rabay, J. and Langendoen, K. (2002) 'Robust positioning algorithms for distributed ad-hoc wireless sensor networks', in *Proc. of the USENIX Technical Annual Conference*.

Savvides, A., Park, H. and Srivastava, M. (2003) 'The n-hop multilateration primitive for node localization problems', *ACM Mobile Networks and Applications*, Vol. 8, pp.443–451.

Sayed, A.H. and Tarighat, A. (2005) 'Network-based wireless location', *IEEE Signal Process. Mag.*, pp.24–40.

Schmidt, R.O. (1972) 'A new approach to geometry of range difference location', *Aerospace and Electronic Systems, IEEE Transactions on*, November, Vol. AES-8, No. 6, pp.821–835.

Tai, Y. and Bo, Y. (2009) 'Collaborative target tracking in wireless sensor network', in *Electronic Measurement and Instruments*, *ICEMI '09*, *9th International Conference on*, August, pp.2–1005–2–1010.

Terwilliger, M., Gupta, A., Bhuse, V., Kamal, Z. and Salahuddin, M. (2004) 'A localization system using wireless network sensors: a comparison of two techniques', in *Proc. of the First Workshop on Positioning, Navigation and Communication*, Hannover, Germany.

World Wide Lightning Location Network (2002) Available at http://webflash.ess.washington.edu.

Yu, K. (2007) '3-d localization error analysis in wireless networks', *IEEE Trans. Wireless Communications*, Vol. 6, No. 10, pp.3473–3481.

On the Quality of Wireless Network Connectivity

Soura Dasgupta
Department of Electrical and Computer Engineering
The University of Iowa

Guoqiang Mao
School of Electrical and Information Engineering
The University of Sydney
National ICT Australia

***Abstract*—Despite intensive research in the area of network connectivity, there is an important category of problems that remain unsolved: how to measure the *quality of connectivity* of a wireless multi-hop network which has a realistic number of nodes, *not necessarily* large enough to warrant the use of asymptotic analysis, and has unreliable connections, reflecting the inherent unreliable characteristics of wireless communications? The quality of connectivity measures how easily and reliably a packet sent by a node can reach another node. It complements the use of *capacity* to measure the quality of a network in saturated traffic scenarios and provides a native measure of the quality of (end-to-end) network connections. In this paper, we explore the use of probabilistic connectivity matrix as a tool to measure the quality of network connectivity. Some interesting properties of the probabilistic connectivity matrix and their connections to the quality of connectivity are demonstrated. We show that the largest eigenvalue of the probabilistic connectivity matrix can serve as a good measure of the quality of network connectivity.**

***Index Terms*—Connectivity, network quality, probabilistic connectivity matrix**

I. Introduction

Connectivity is one of the most fundamental properties of wireless multi-hop networks [1]–[3], and is a prerequisite for providing many network functions. A network is said to be *connected* if and only if (iff) there is a (multi-hop) path between any pair of nodes. Further, a network is said to be k-connected iff there are k mutually independent paths between any pair of nodes that do not share any node in common except the starting and the ending nodes. k-connectivity is often required for robust operations of the network.

There are two general approaches to studying the connectivity problem. The first, spearheaded by the seminal work of Penrose [3] and Gupta and Kumar [1], is based on an asymptotic analysis of large-scale random networks, which considers a network of n nodes that are *i.i.d.* on an area with an underlying uniform distribution. A pair of nodes are directly connected iff their Euclidean distance is smaller than or equal to a given threshold $r(n)$, independent of other connections. Some interesting results are obtained on the value of $r(n)$ required for the above network to be *asymptotically almost surely* connected as $n \to \infty$. In [4], [5], the authors extended the above results from the unit disk model to a random connection model, in which any pair of nodes separated by a displacement $\boldsymbol{x}$ are directly connected with probability $g(\boldsymbol{x})$, independent of other connections. We refer readers to [7] for a more comprehensive review of related work.

The second approach is based on a deterministic setting and studies the connectivity and other topological properties of a network using algebraic graph theory. Specifically, consider a network with a set of n nodes. Its property can be studied using its *underlying graph* $G(V,E)$, where $V \triangleq \{v_1,\ldots,v_n\}$ denotes the vertex set and E denotes the edge set. The underlying graph is obtained by representing each node in the network uniquely using a vertex and the converse. An undirected edge exists between two vertices iff there is a direct connection (or link) between the associated nodes[1]. Define an *adjacency matrix* A_G of the graph $G(V,E)$ to be a symmetric $n \times n$ matrix whose $(i,j)^{th}, i \neq j$ entry is equal to one if there is an edge between v_i and v_j and is equal to zero otherwise. Further, the diagonal entries of A_G are all equal to zero. The *eigenvalues of the graph* $G(V,E)$ are defined to be the eigenvalues of A_G. The network connectivity information, e.g. connectivity and k-connectivity, is entirely contained in its adjacency matrix. Many interesting connectivity and topological properties of the network can be obtained by investigating the eigenvalues of its underlying graph. For example, let $\mu_1 \geq \ldots \geq \mu_n$ be the eigenvalues of a graph G. If $\mu_1 = \mu_2$, then G is disconnected. If $\mu_1 = -\mu_n$ and G is not empty, then at least one connected component of G is nonempty and bipartite. If the number of distinct eigenvalues of G is r, then G has a *diameter* of at most $r-1$ [8]. Some researchers have also studied the properties of the underlying graph using its Laplacian matrix [9], where the Laplacian matrix of a graph G is defined as $L_G \triangleq D - A_G$ and D is a diagonal matrix with degrees of vertices in G on the diagonal. Particularly, the *algebraic connectivity* of a graph G is the second-smallest eigenvalue of L_G and it is greater than 0 iff G is a connected graph. The algebraic connectivity quantifies the speed of convergence of consensus algorithms [10]. We refer readers to [8] for a comprehensive treatment of the topic.

Despite intensive research in the area, there is an important category of problems that remain unsolved: how to measure the *quality of connectivity* of a wireless multi-hop

This research is partially supported by US NSF grants ECS-0622017, CCF-072902, and CCF-0830747.

This research is partially supported by ARC Discovery project DP110100538 and by the Air Force Research Laboratory, under agreement number FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

[1]In this paper, we limit our discussions to a *simple graph* (network) where there is at most one edge (link) between a pair of vertices (nodes) and an undirected graph.

network which has a realistic number of nodes, *not necessarily* large enough to warrant the use of asymptotic analysis, and has unreliable connections, reflecting the inherent unreliable characteristics of wireless communications? The quality of connectivity measures how easily and reliably a packet sent by a node can reach another node. It complements the use of *capacity* to measure the quality of a network in saturated traffic scenarios and provides a native measure of the quality of (end-to-end) network connections. In the following paragraphs, we elaborate on the above question using two examples.

Example 1: Consider a network with a fixed number of nodes with known transmission power to be deployed in a region. Assume that the wireless propagation model in that environment is known and its characteristics have been quantified through *a priori* measurements or empirical estimation. Further, a link exists between two nodes iff the received signal strength from one node at the other node is greater than or equal to a predetermined threshold *and* the same is also true in the opposite direction. One can then find the probability that a link exists between two nodes at two fixed locations: It is determined by the probability that the received signal strength is greater than or equal to the pre-determined threshold. Two related questions can be asked: a) If these nodes are deployed at a set of known locations, what is the quality of connectivity of the network, measured by the probability that there is a path between any two nodes, as compared to node deployment at another set of locations? b) How to optimize the node deployment to maximize the quality of connectivity?

Example 2: Consider a network with a fixed number of nodes. The transmission between a pair of nodes with a direct connection quantifying the inherent unreliable characteristics of wireless communications. There are no direct connections between some pairs of nodes because the probability of successful transmission between them is too low to be acceptable. How to measure the quality of connectivity of such a network, in the sense that a packet transmitted from one node can easily and reliably reach another node via a multi-hop path. Will a single "good" path between a pair of nodes be more preferable than multiple "bad" paths? These are further illustrated using Fig. 1 and 2.

In this paper, we explore the use of probabilistic connectivity matrix, a concept to be defined later in Section II, as a tool to measure the quality of network connectivity. Some interesting properties of the probabilistic connectivity matrix and their connections to the quality of connectivity are demonstrated. Based on the analysis, we show that the largest eigenvalue of the probabilistic connectivity matrix can serve as a good metric of the quality of network connectivity.

The rest of the paper is organized as follows. Section II defines the network settings, the probabilistic connectivity matrix and gives a method to compute the matrix. Section III introduces certain inequalities associated with the entries of the probabilistic connectivity matrix. Section IV proves several important results about the probabilistic connectivity matrix. These directly associate the largest eigenvalue of the probabilistic connectivity matrix to the quality of connectivity and expose a structure that holds the promise of facilitating associated optimization tasks. Section V concludes the paper and discusses future work.

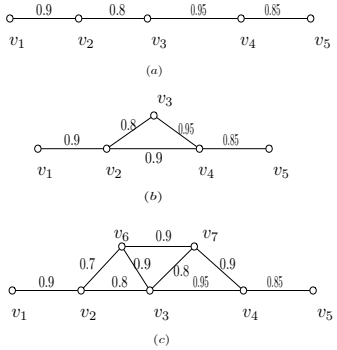


Figure 1: An illustration of networks with different quality of connectivity. A solid line represents a direct connection between two nodes and the number beside the line represents the corresponding transmission successful probability. The networks shown in (a), (b), and (c) are all connected networks but not 2-connected networks, i.e. their connectivity cannot be differentiated using the k-connectivity concept. However intuitively the quality of the network in (b) is better than that of the network in (a) because of the availability of the additional high-quality link between v_2 and v_4 in (b). The quality of the network in (c) is even better because of the availability of the additional nodes and the associated high-quality links, hence additional routes, if these additional nodes act as relay nodes only. If these additional nodes also generate their own traffic, it is uncertain whether the quality of the network in (c) is better or not. Therefore it is important to develop a measure to quantitatively compare the quality of connectivity (for the networks in (a) and (b)) and to evaluate the benefit of additional nodes on connectivity (for the network in (c)).

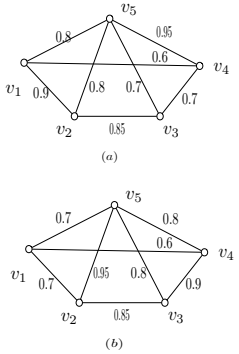


Figure 2: The networks shown in (a) and (b) have the same topology but different link quality. It is difficult to compare the quality of the two networks.

II. Definition and Construction of the Probabilistic Connectivity Matrix

Consider a network of n nodes. For some pair of nodes, an edge (or link) may exist with a non-negligible probability. The edges are undirected and independent.

Denote the underlying graph of the above network by $G(V, E)$, where $V = \{v_1, \ldots, v_n\}$ is the vertex set and $E = \{e_1, \ldots, e_m\}$ is the edge set, which contains the set of

all possible edges. Here the vertices and the edges are indexed from 1 to n and from 1 to m respectively. For convenience, in some parts of this paper we also use the symbol e_{ij} to denote an edge between vertices v_i and v_j when there is no confusion. We associate with each edge e_i, $i \in \{1, \dots m\}$, an indicator random variable I_i such that $I_i = 1$ if the edge e_i exists; $I_i = 0$ if the edge e_i does not exist. The indicator random variables I_{ij}, $i \neq j$ and $i, j \in \{1, \dots n\}$, are defined analogously.

In the following, we give a definition of the probabilistic adjacency matrix:

Definition 1: The probabilistic adjacency matrix of $G(V, E)$, denoted by A_G, is a $n \times n$ matrix such that its $(i,j)^{th}$, $i \neq j$, entry $a_{ij} \triangleq \Pr(I_{ij} = 1)$ and its diagonal entries are all equal to 1.

Due to the undirected property of an edge mentioned above, A_G is a symmetric matrix, i.e. $a_{ij} = a_{ji}$. Note that the diagonal entries of A_G are defined to be 1, which is different from that common in the literature. This treatment of the diagonal entries can be associated with the fact that a node in the network can store a packet until better transmission opportunity arises when it finds the wireless channel busy [11].

The probabilistic connectivity matrix is defined in the following:

Definition 2: The probabilistic connectivity matrix of $G(V, E)$, denoted by Q_G, is a $n \times n$ matrix such that its $(i,j)^{th}$, $i \neq j$, entry is the probability that there exists a path between vertices v_i and v_j, and its diagonal entries are all equal to 1.

As a ready consequence of the symmetry of A_G, Q_G is also a symmetric matrix.

Given the probabilistic adjacency matrix A_G, the probabilistic connectivity matrix Q_G is fully determined. However the computation of Q_G is not trivial because for a pair of vertices v_i and v_j, there may be multiple paths between them and some of them may share common edges, i.e. are not *independent*. In the following paragraph, we give an approach to computing the probabilistic connectivity matrix.

Let $(I_1, \dots, I_m)$ be a particular instance of the indicator random variables associated with an instance of the random edge set. Let $Q_G|(I_1, \dots, I_m)$ be the connectivity matrix of G *conditioned on* $(I_1, \dots, I_m)$. The $(i,j)^{th}$ entry of $Q_G|(I_1, \dots, I_m)$ is either 0, when there is no path between v_i and v_j, or 1 when there exists a path between v_i and v_j. The diagonal entries of $Q_G|(I_1, \dots, I_m)$ are always 1. Conditioned on $(I_1, \dots, I_m)$, $G(V, E)$ is just a deterministic graph. Therefore the entries of $Q_G|(I_1, \dots, I_m)$ can be efficiently computed using a search algorithm, such as breadth-first search. Given $Q_G|(I_1, \dots, I_m)$, Q_G can be computed using the following equation:

$$Q_G = E\left(Q_G|(I_1, \dots, I_m)\right) \tag{1}$$

where the expectation is taken over all possible instances of $(I_1, \dots, I_m)$.

The approach suggested in the last paragraph is essentially a brute-force approach to computing Q_G. A more efficient algorithm is suggested in Section IV.

Remark 1: A major difference between the (probabilistic) connectivity matrix and the adjacency matrix (or the Laplacian matrix) is that the later matrix focuses on quantifying the relation between node pairs directly connected by an edge only while the former matrix focuses on quantifying the end-to-end relationship between node pairs. It is not trivial to obtain the connectivity matrix from the adjacency matrix or use the adjacency matrix to study network properties easily obtainable using the connectivity matrix.

Remark 2: For simplicity, the terms used in our discussion are based on the problems in Example 1. The discussion however can be easily adapted to the analysis of the problems in Example 2. For example, if a_{ij} is defined to be the probability that a transmission between nodes v_i and v_j is successful, the $(i,j)^{th}$ entry of the probabilistic connectivity matrix Q_G computed using (1) then gives the probability that a transmission from v_i to v_j via a multi-hop path is successful under the best routing algorithm, which can always find a shortest and error-free path from v_i to v_j if it exists, or alternatively, the probability that a packet broadcast from v_i can reach v_j where each node receiving the packet only broadcasts the packet once. Therefore the $(i,j)^{th}$ entry of Q_G can be used as a quality measure of the end-to-end paths between v_i and v_j, which takes into account the fact that availability of extra paths between a pair of nodes can be exploited to improve the probability of successful transmissions.

III. Some Key Inequalities for Connection Probabilities

The entries of the probabilistic connectivity matrix give a measure of the quality of end-to-end paths. In this section, we provide some important inequalities that may facilitate further analysis of the quality of connectivity. Some of these inequalities are exploited in the next section to establish several key properties of the probabilistic connection matrix itself. We first introduce some results that are required for the further analysis of the probabilistic connectivity matrix Q_G.

For a random graph with a given set of vertices, a particular event is *increasing* if the event is preserved when more edges are added into the graph. An event is *decreasing* if its complement is increasing.

Denote by ξ_{ij} the event that there is a path between vertices v_i and v_j, $i \neq j$. Denote by ξ_{ikj} the event that there is a path between vertices v_i and v_j *and* that path passes through the third vertex v_k, where $k \in \Gamma_n \backslash \{i, j\}$ and Γ_n is the set of indices of all vertices. Denote by η_{ij} the event that there is an edge between vertices v_i and v_j. Denote by π_{ikj} the event that there is a path between vertices v_i and v_k *and* there is a path between vertices v_k and v_j, where $k \in \Gamma_n \backslash \{i, j\}$. It can be shown from the above definitions that

$$\xi_{ij} = \eta_{ij} \cup \left(\cup_{k \neq i,j} \xi_{ikj}\right) \tag{2}$$

Let q_{ij}, $i \neq j$, be the $(i,j)^{th}$ entry of Q_G, i.e. $q_{ij} = \Pr(\xi_{ij})$. The following lemma can be readily obtained from the FKG inequality [6, Theorem 1.4] and the above definitions

Lemma 1: For two distinct indices $i, j \in \Gamma_n$ and $\forall k \in \Gamma_n \backslash \{i, j\}$

$$q_{ij} \geq \max_{k \in \Gamma_n \backslash \{i,j\}} q_{ik} q_{kj} \tag{3}$$

Proof: It follows readily from the above definitions that the event ξ_{ij} is an increasing event. Using the FKG inequality:

$$\Pr(\xi_{ij}) \geq \Pr(\pi_{ikj}) = \Pr(\xi_{ik} \cap \xi_{kj}) \geq \Pr(\xi_{ik}) \Pr(\xi_{kj}) \tag{4}$$

■

Lemma 1 gives a lower bound of q_{ij}. The following lemma gives an upper bound of q_{ij}:

Lemma 2: For two distinct indices $i, j \in \Gamma_n$ and $\forall k \in \Gamma_n \backslash \{i, j\}$,

$$q_{ij} \leq 1 - (1 - a_{ij}) \prod_{k \in \Gamma_n \backslash \{i,j\}} (1 - q_{ik} q_{kj}) \tag{5}$$

where $a_{ij} = \Pr(\eta_{ij})$.

Proof: We will first show that $\xi_{ikj} \Leftrightarrow \xi_{ik} \square \xi_{kj}$. That is, the occurrence of the event ξ_{ikj} is a sufficient and necessary condition for the occurrence of the event $\xi_{ik} \square \xi_{kj}$, where for two events A and B, $A \square B$ denotes the event that there exist two *disjoint* sets of edges such that the first set of edges guarantees the occurrence of A and the second set of edges guarantees the occurrence of B.

Using the definition of ξ_{ikj}, occurrence of ξ_{ikj} means that there is a path between vertices v_i and v_j *and* that path passes through vertex v_k. It follows that there exist a path between vertex i and vertex v_k *and* a path between vertex v_k and vertex v_j *and* the two paths do not have edge(s) in common. Otherwise, it will contradict the definition of ξ_{ikj}, particularly as the definition of a path requires the edges to be distinct. Therefore $\xi_{ikj} \Rightarrow \xi_{ik} \square \xi_{kj}$. Likewise, $\xi_{ikj} \Leftarrow \xi_{ik} \square \xi_{kj}$ also follows directly from the definitions of ξ_{ikj}, ξ_{ik}, ξ_{kj} and $\xi_{ik} \square \xi_{kj}$. Consequently

$$\Pr(\xi_{ikj}) = \Pr(\xi_{ik} \square \xi_{kj}) \leq \Pr(\xi_{ik}) \Pr(\xi_{kj}) \tag{6}$$

where the inequality is a direct result of the BK inequality [6]

.

With a little bit abuse of the terminology, in the following derivations we also use ξ_{ikj} to represent the set of edges that make the event ξ_{ikj} happen, and use η_{ij} to denote the edge between vertices v_i and v_j.

Note that the set of edges $\cup_{k \in \Gamma_n \backslash \{i,j\}} \xi_{ikj}$ does not contain η_{ij}. Therefore using (2) and independence of edges (used in the third step)

$$\begin{aligned}
q_{ij} &= \Pr\left(\eta_{ij} \cup \left(\cup_{k \in \Gamma_n \backslash \{i,j\}} \xi_{ikj}\right)\right) \\
&= 1 - \Pr\left(\overline{\eta_{ij}} \cap \left(\overline{\cup_{k \in \Gamma_n \backslash \{i,j\}} \xi_{ikj}}\right)\right) \\
&= 1 - (1 - a_{ij}) \Pr\left(\cap_{k \in \Gamma_n \backslash \{i,j\}} \overline{\xi_{ikj}}\right) \\
&\leq 1 - (1 - a_{ij}) \prod_{k \in \Gamma_n \backslash \{i,j\}} \Pr\left(\overline{\xi_{ikj}}\right) \qquad (7) \\
&= 1 - (1 - a_{ij}) \prod_{k \in \Gamma_n \backslash \{i,j\}} (1 - \Pr(\xi_{ikj})) \\
&\leq 1 - (1 - a_{ij}) \prod_{k \in \Gamma_n \backslash \{i,j\}} (1 - q_{ik} q_{kj}) \qquad (8)
\end{aligned}$$

where in (7), FKG inequality and the fact that $\overline{\xi_{ikj}}$ is a decreasing event are used and the last step results due to (6).

■

When there is no edge between vertices v_i and v_j, which is the generic case, the upper and lower bounds in Lemmas 1 and 2 reduce to

$$\max_{k \in \Gamma_n \backslash \{i,j\}} q_{ik} q_{kj} \leq q_{ij} \leq 1 - \prod_{k \in \Gamma_n \backslash \{i,j\}} (1 - q_{ik} q_{kj}) \tag{9}$$

The above inequality sheds insight on how the quality of paths between a pair of vertices is related to the quality of paths between other pairs of vertices. It can be possibly used to determine the most effective way of improving the quality of a particular set of paths by improving the quality of a particular (set of) edge(s), or equivalently what can be reasonably expected from an improvement of a particular edge on the quality of end-to-end paths.

The following lemma further shows that relation among entries of the path matrix Q_G can be further used to derive some topological information of the graph.

Lemma 3: If $q_{ij} = q_{ik} q_{kj}$ for three distinct vertices v_i, v_j and v_k, the vertex set V of the underlying graph $G(V, E)$ can be divided into three non-empty and non-intersecting sub-sets V_1, V_2 and V_3 such that $v_i \in V_1$, $v_j \in V_3$ and $V_2 = \{v_k\}$ *and* any possible path between a vertex in V_1 and a vertex in V_2 must pass through v_k. Further, for any pair of vertices v_l and v_m, where $v_l \in V_1$ and $v_m \in V_3$, $q_{lm} = q_{lk} q_{km}$.

Proof: Using (4) in the second step, it follows that

$$\begin{aligned}
q_{ij} = \Pr((\xi_{ij} \backslash \pi_{ikj}) \cup \pi_{ikj}) &= \Pr(\xi_{ij} \backslash \pi_{ikj}) + \Pr(\pi_{ikj}) \\
&\geq \Pr(\xi_{ij} \backslash \xi_{ikj}) + q_{ik} q_{kj}
\end{aligned}$$

Therefore $q_{ij} = q_{ik} q_{kj}$ implies that $\Pr(\xi_{ij} \backslash \pi_{ikj}) = 0$ or equivalently $\xi_{ij} \Leftrightarrow \pi_{ikj}$

Further, $\Pr(\xi_{ij} \backslash \pi_{ikj}) = 0$ implies that a *possible* path (i.e. a path with a non-zero probability) connecting v_i and v_k and a *possible* path connecting v_k and v_j cannot have any edge in common. Otherwise a path from v_i to v_j, bypassing v_k, exists with a non-zero probability which implies $\Pr(\xi_{ij} \backslash \xi_{ikj}) > 0$. The conclusion follows readily that if $q_{ij} = q_{ik} q_{kj}$ for three distinct vertices v_i, v_j and v_k, the vertex set V of the underlying graph $G(V, E)$ can be divided into three non-empty and non-overlapping sub-sets V_1, V_2 and V_3 such that $v_i \in V_1$, $v_j \in V_3$ and $V_2 = \{v_k\}$ *and* a path between a vertex in V_1 and a vertex in V_2, if exists, must pass through v_k.

Further, for any pair of vertices v_l and v_m, where $v_l \in V_1$ and $v_m \in V_3$, it is easily shown that $\Pr(\xi_{lm} \backslash \pi_{lkm}) = 0$. Due to independence of edges and further using the fact that $\Pr(\xi_{lm} \backslash \pi_{lkm}) = 0$, it can be shown that

$$\Pr(\xi_{lm}) = \Pr(\pi_{lkm}) = \Pr(\xi_{lk} \cap \xi_{km}) = \Pr(\xi_{lk}) \Pr(\xi_{km})$$

where the last step results because under the condition of $\Pr(\xi_{lm} \backslash \pi_{lkm}) = 0$, a path between v_l and v_k and a path between v_k and v_m cannot possibly have any edge in common.

■

An implication of Lemma 3 is that for any three distinct vertices, v_i, v_j and v_k, if a relationship $q_{ij} = q_{ik} q_{kj}$ holds, vertex v_k must be a *critical* vertex whose removal will render the graph disconnected.

IV. Properties of the Connectivity Matrix

Having established some inequalities obeyed by the entries of Q_G, we now turn to establishing a measure of the quality of network connectivity. At the core of the development in this section is the following result.

Lemma 4: Each off-diagonal entry of the probabilistic connectivity matrix Q_G is a multiaffine[2] function of a_{ij}, $i \in \{1,\ldots,n\}, j > i$.

Proof: Observe that $a_{ij} = \Pr(\eta_{ij})$ and the events η_{ij}, $i \in \{1,\ldots,n\}, j > i$ are independent. The conclusion in the lemma follows readily from the fact that the event associated with each q_{ij}, i.e. there exists a path between vertices v_i and v_j, is a union of intersections of these events η_{ij}, $i \in \{1,\ldots,n\}, j > i$. ∎

Due to the above multiaffine property, for any four positive integers $k,l,i,j \in \{1,\ldots n\}$, where $p \neq q$ and $i \neq j$, the following holds:

$$q_{lk} = f(E\backslash\{e_{ij}\})\, a_{ij} + g(E\backslash\{e_{ij}\}) \tag{10}$$

where $f(E\backslash\{e_{ij}\})$ and $g(E\backslash\{e_{ij}\})$ are non-negative constants within $[0,1]$ determined by the state of the set of edges *excluding* e_{ij}. $g(E\backslash\{e_{ij}\}) = 0$ implies that non-existence of the edge e_{ij} will render the vertices v_l and v_k disconnected. $f(E\backslash\{e_{ij}\}) = 0$ implies that the state of the edge e_{ij} is irrelevant for the end-to-end paths between v_l and v_k. Further, $f(E\backslash\{e_{ij}\})$ can be used to measure the *criticality* of the edge e_{ij} to the end-to-end paths between v_l and v_k.

Remark 3: Using the multiaffine property, a more efficient algorithm for computing Q_G than the one suggested earlier using (1) can be constructed. Particularly, the probabilistic connectivity matrix of a network forming a tree can be easily computed. Therefore the algorithm may start by first identifying a spanning tree in $G(V,E)$ and computing the associated probabilistic connectivity matrix. Then, the edges in E but outside the spanning tree can be added recursively and the corresponding probabilistic connectivity matrix updated using (10).

We comment later in Remark 5 on how the multiaffine structure is also potentially useful for performing some of the optimization tasks inherent in maximizing connectivity. e.g. determination of the link whose improvement will bring the maximum benefit on connectivity.

A very desirable property of Q_G is established below.

Theorem 1: The probabilistic connectivity matrix Q_G is a positive semi-definite matrix. Further, Q_G is positive semi-definite but not positive definite iff there exist distinct $i,j \in \{1,\cdots,n\}$, such that $q_{ij} = 1$.

The proof is omitted due to space limitation.

Let $\lambda_1 \geq \ldots \geq \lambda_n$ be the eigenvalues of Q_G. Note that $\lambda_1 + \cdots + \lambda_n = n$. As an easy consequence of Theorem 1, $n \geq \lambda_1 \geq 1$ and $1 \geq \lambda_n \geq 0$. In the best case, Q_G is a matrix with all entries equal to 1. Then $\lambda_1 = n$ and $\lambda_2 = \cdots = \lambda_n = 0$. In the worst case, Q_G is an identity matrix. Then $\lambda_1 = \cdots = \lambda_n = 1$. This suggests that λ_1, i.e. the largest eigenvalue of Q_G, can be used as a measure of quality of network connectivity and a larger λ_1 indicates a better quality.

Further, let X be a vector representing the number of packets broadcast by each node to the rest of the network and let Y be a vector representing the *random* number of packets received by each node. It is obvious that $E[Y|X] = Q_G X$ then represents the expected number of packets received by each node. Using the property that Q_G is a symmetric matrix, it can be shown that

$$\begin{aligned} \max_{||X||_2=1} ||E[Y|X]||_2 &= \max_{||X||_2=1} ||Q_G X||_2 \\ = \max_{||X||_2=1} \sqrt{X^T Q_G^T Q_G X} &= \max_{||X||_2=1} \sqrt{X^T Q_G^2 X} \\ = \sqrt{\lambda_{max}(Q_G^2)} = \sqrt{\lambda_{max}^2(Q_G)} &= \lambda_{max}(Q_G) \end{aligned}$$

where $\lambda_{\max}(Q_G)$ is the maximum eigenvalue of Q_G and $||X||_2$ denotes the L^2-norm or Euclidean norm of X.

We will make this idea that $\lambda_{\max}(Q_G)$ serves as a good measure of the quality of network connectivity more concrete in the following analysis. We start our discussion with a connected network and then extend to more generic cases. We will call a network *connected* if for all $i,j \in \{1,\cdots,n\}$, $q_{ij} > 0$. Obviously the probabilistic connectivity matrix of a connected network is *irreducible* [12, p. 374] as all the entries of the matrix are non-zero. As a measure of the quality of network connectivity, if the path probabilities q_{ij} increase, the largest eigenvalue of the probabilistic connectivity matrix should also increase. This is formally stated below:

Theorem 2: Let $G(V,E)$ and $G'(V,E')$ be the underlying graphs of two connected networks defined on the same vertex set V but with different link probabilities. Let Q_G and $Q_{G'}$ be the probabilistic connectivity matrices of G and G' respectively. If $Q'_G - Q_G$ is a non-zero, non-negative matrix[3], then $\lambda_{\max}(Q_G) < \lambda_{\max}(Q_{G'})$.

Proof:

We need the following lemma to prove Theorem 2.

Lemma 5: Suppose $A = A^T \neq B = B^T$ are non-negative, irreducible, real matrices, and $B - A$ is a non-zero, non-negative matrix. Then: $\lambda_{\max}(A) < \lambda_{\max}(B)$.

Proof: Observe at least one element of $B - A$ is positive. From Perron-Frobenius theorem [12, p. 536], $x \in \mathbb{R}^n$, the eigenvector corresponding to the largest eigenvalue of A can be chosen to have all elements positive. Then the result follows from the fact that:

$$\begin{aligned} \lambda_{\max}(A) x^T x &= x^T A x \\ &= x^T B x - x^T (B - A) x \\ &< x^T B x \leq \lambda_{\max}(B) x^T x \end{aligned}$$

as $B - A$ is a non-zero, non-negative matrix. ∎

Turning to the proof of Theorem 2 we note that the result follows directly from Lemma 5 and the fact $Q_{G'}$ and Q_G satisfy the requirements of B and A, respectively.

∎

If the network is not connected, i.e. some entries of its probabilistic connectivity matrix is 0, the network can be

[2] A multiaffine function is affine in each variable when the other variables are fixed.

[3] A matrix is *non-negative* if all its entries are greater than or equal to 0.

decomposed into *disjoint components*. Let the total number of components in the network be k. Let G_i be the subgraph induced on the set of vertices in the i^{th} component and Q_{G_i} be the probabilistic connectivity matrix of G_i. It follows that

$$\lambda_{\max}(Q_G) = \max\{\lambda_{\max}(Q_{G_1}), \cdots, \lambda_{\max}(Q_{G_k})\} \quad (11)$$

We consider two basic situations: a) there are increases in some entries of Q_G from non-zero values but such increases do not change the number of components in the network. It then follows easily from Theorem 2 that $\lambda_{\max}\left(Q_{G_i'}\right) > \lambda_{\max}(Q_{G_i})$. Depending on whether $\lambda_{\max}\left(Q_{G_i'}\right)$ is greater than $\lambda_{\max}(Q_G)$ or not however, $\lambda_{\max}(Q_G)$ may or may not increase. b) there are increases in some entries of Q_G from zero to non-zero values and such increases reduce the number of components in the network. For situation b), we consider a simplified scenario where increases in the path probabilities merge two originally disjoint components, denoted by G_i and G_j. The more complicated scenario where increases in the path probabilities join more than two originally disjoint components can be obtained recursively as an extension of the above simplified scenario. Let G' be the underlying graph of the network after increases in path probabilities and let G'_{ij} be the subgraph in G' induced on the vertex set $V_i \cup V_j$. Obviously $Q_{G'_{ij}}$ is an irreducible matrix and the following result can be established.

Lemma 6: Under the above settings,

$$\lambda_{\max}\left(Q_{G'_{ij}}\right) > \lambda_{\max}\left(\text{ diag }\{Q_{G_i}, Q_{G_j}\}\right) \quad (12)$$

The proof of Lemma 6 is straightforward and hence omitted.

Thus indeed the largest eigenvalues of the probabilistic connection matrices associated with disjoint components measure the quality of the components connection.

Remark 4: To compare two networks with different number of nodes, the normalized maximum eigenvalue of the probabilistic connectivity matrix, where the maximum eigenvalue is divided by the number of nodes, can be used.

Remark 5: The fact that the largest eigenvalue of the probabilistic connectivity matrix measures connectivity, suggests the following obvious optimization. Modify one or more a_{ij} under suitable constraints to maximize the largest eigenvalue of the probabilistic connectivity matrix. Results in [13] and [14] suggest that the multiaffine dependence of the q_{ij} on the a_{ij} *together with the fact that* Q_G *is positive semi-definite* promise to facilitate such optimization.

V. Conclusions and Further Work

In this paper we explored the use of the probabilistic connectivity matrix as a tool to measure the quality of network connectivity. Some interesting properties of the probabilistic connectivity matrix and their connections to the quality of network connectivity were demonstrated. Particularly, the off-diagonal entries of the probabilistic connectivity matrix provide a measure of the quality of end-to-end connections and we have also provided theoretical analysis supporting the use of the largest eigenvalue of the probabilistic connectivity matrix as a measure of the quality of overall network connectivity.

Inequalities between the entries of the probabilistic connectivity matrix were established. These may provide insights into the correlations between quality of end-to-end connections. Further, the probabilistic connectivity matrix was shown to be a positive semi-definite matrix and its off-diagonal entries are multiaffine functions of link probabilities. These two properties are expected to be very helpful in optimization and robust network design, e.g. determining the link whose quality improvement will result in the maximum gain in network quality, and determining quantitatively the relative criticality of a link to either a particular end-to-end connection or to the entire network.

The results in the paper rely on two main assumptions: the links are symmetric *and* independent. We expect that our analysis can be readily extended such that the first assumption on symmetric links can be removed – in fact the results in Section III do not need this assumption. While in the asymmetric case the probabilistic connectivity matrix is no longer guaranteed to be positive semi-definite, we conjecture that the largest eigenvalue retains its significance. Discarding the second assumption requires more work. However, we are encouraged by the following observation. If we introduce conditional edge probabilities into the mix, then Q_G is still a multiaffine function of the a_{ij} and the conditional probabilities. Thus we still expect all the results in Section IV to hold, though the proof may be non-trivial. In real applications link correlations may arise due to both physical layer correlations and correlations caused by traffic congestion.

References

[1] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," in *Stochastic Analysis, Control, Optimization and Applications*. Boston, MA: Birkhauser, 1998, pp. 547–566.

[2] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE JSAC*, vol. 27, no. 7, pp. 1029–1046, 2009.

[3] M. D. Penrose, *Random Geometric Graphs*, Oxford University Press, 2003.

[4] G. Mao and B. D. Anderson, "Connectivity of large scale networks: Emergence of unique unbounded component," *submitted to IEEE TMC, available at http://arxiv.org/abs/1103.1991*, 2011.

[5] ——, "Connectivity of large scale networks: Distribution of isolated nodes," *submitted to IEEE TMC, available at http://arxiv.org/abs/1103.1994*, 2011.

[6] R. Meester and R. Roy, *Continuum Percolation*, Cambridge University Press, 1996.

[7] G. Mao and B. D. Anderson, "Towards a better understanding of large scale network models," *accepted to appear in IEEE/ACM ToN*, 2011.

[8] N. L. Biggs, *Algebraic Graph Theory*. Cambridge University Press, 1974.

[9] B. Mohar, "Laplace eigenvalues of graphs - a survey," *Journal of Discrete Mathematics*, vol. 109, pp. 171–183, 1992.

[10] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[11] G. Mao and B. D. Anderson, "Graph theoretic models and tools for the analysis of dynamic wireless multihop networks," in *IEEE WCNC*, 2009, pp. 1–6.

[12] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, ser. Computer Science and Applied Mathematics. Academic Press, 1985.

[13] L. Zadeh and C. A. . Desoer, *Linear Systems Theory*. McGraw Hill, New York, 1963.

[14] S. Dasgupta, C. Chockalingam, M. Fu, and B. Anderson, "Lyapunov functions for uncertain systems with applications to the stability of time varying systems," *IEEE Transactions on Circuits and Systems-I, Fundamental Theory and Applications*, vol. 41, no. 2, pp. 93–105, 1994.

Wireless Multi-hop Networks: Current Research and Future Challenges

Guoqiang Mao
School of Electrical and Information Engineering
The University of Sydney
National ICT Australia
Email: guoqiang.mao@sydney.edu.au

***Abstract*— Wireless multi-hop networks, in various forms and under various names, are being increasingly used in military and civilian applications. Studying connectivity and capacity of these networks is an important problem. The scaling behavior of connectivity and capacity when the network becomes sufficiently large is of particular interest. In this paper, we briefly overview recent development and discuss research challenges and opportunities in the area, with a focus on the network connectivity. We demonstrate some intrinsic connections between the connectivity analysis and capacity analysis and point out the fundamental parameters determining the capacity of a wireless multi-hop network.**

***Index Terms*— Wireless multi-hop networks, connectivity, capacity**

I. Introduction

Wireless multi-hop networks, in various forms, e.g. wireless sensor networks, underwater sensor networks, vehicular networks, mesh networks and UAV (Unmanned Aerial Vehicle) formations, and under various names, e.g. ad-hoc networks, hybrid networks, delay tolerant networks and intermittently connected networks, are being increasingly used in military and civilian applications. There are three defining features that characterize a wireless multi-hop network:

1) Wireless devices are self-organized or assisted by some infrastructure to form a network. The former case corresponds to ad-hoc networks whereas the latter case corresponds to infrastructure-based multi-hop networks. Depending on the applications, the forms of the infrastructure can be quite flexible, e.g. a subset of devices connected via wired connections, a subset of devices with more powerful transmission capability such that they form a wireless backbone for the network, or in a UAV formation, the infrastructure may assume the form of a subset of UAVs with satellite links.
2) Communication is mostly via wireless multi-hop paths.
3) Packets are forwarded *collaboratively* from the source to the destination.

The implication of the first feature is that ad-hoc networks form an important special case of wireless multi-hop networks but the concept of wireless multi-hop networks has much broader meaning. Particularly, the prospect of including infrastructure into the ad-hoc network addresses shortcomings of the ad-hoc network in scalability and providing reliable service. The second feature sets wireless multi-hop networks apart from the traditional one-hop networks, i.e. cellular networks and wireless LANs. Therefore, there is a unique set of challenging problems specific to wireless multi-hop networks. The third feature implies that collaborative communication, either centrally designed and operated or performed distributedly, is an important consideration in wireless multi-hop networks.

Studying connectivity and capacity of wireless multi-hop networks is an important problem [1]–[3]. The scaling behavior of connectivity and capacity when the network becomes sufficiently large is of particular interest. In this paper, we briefly overview recent development and discuss research challenges and opportunities in the area, with a focus on the network connectivity. We also demonstrate some intrinsic connections between the connectivity results and capacity results and point out the fundamental parameters determining the capacity of a wireless multi-hop network.

A network is said to be *connected* iff (if and only if) there is a (multi-hop) path between any pair of nodes. Further, a network is said to be k-connected if there are k mutually independent paths between any pair of nodes that do not have any node in common except the starting and the ending nodes. k-connectivity is often required for robust operations of the network.

The rest of the paper is organized as follows: Section II discusses connectivity of large-scale random networks; Section III discusses connectivity of giant component; Section IV discusses recent development, research chal-

Manuscript received February 21, 2012; revised March 28, 2012; accepted March 28, 2012.

This research is partially supported by ARC Discovery projects DP110100538 and DP110100538. This material is based on research partially sponsored by the Air Force Research Laboratory, under agreement number FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

This paper is an extended version of an invited position paper published in International Conference on Computing, Networking and Communications 2012.

lenges and opportunities in mobile networks and Section V concludes the paper.

II. Connectivity of Large-Scale Random Networks

A. Unit disk model and connectivity

Extensive research has been done on connectivity problems using the well-known *random geometric graph* and the *unit disk model*, which is usually obtained by randomly and uniformly distributing n nodes in a given area and connecting any two nodes iff their Euclidean distance is smaller than or equal to a given threshold $r(n)$, known as the *transmission range* [3], [4]. Significant outcomes have been achieved for both asymptotically infinite n [1], [3], [5]–[10] and finite n [11]–[14].

Research on the connectivity of large-scale random ad-hoc networks under the unit disk model is spearheaded by Penrose [15], [16] and Gupta and Kumar [1]. Specifically, Penrose [15], [16] and Gupta and Kumar [1] proved using different techniques that if the transmission range is set to $r(n) = \sqrt{\frac{\log n + c(n)}{\pi n}}$, a random network formed by uniformly placing n nodes on a unit-area disk in $\Re^2$ is *asymptotically almost surely* (a.a.s.) connected as $n \to \infty$ iff $c(n) \to \infty$. An event ξ_n depending on n is said to occur a.a.s. if its probability tends to one as $n \to \infty$. Penrose's result is based on the fact that in the above random network, as $n \to \infty$ the longest edge of the minimum spanning tree converges in probability to the minimum transmission range required for the above random network to have no isolated nodes (or equivalently the longest edge of the nearest neighbor graph of the above network) [3], [15], [16]. Gupta and Kumar's result is based on a key finding in the continuum percolation theory [17, Chapter 6]: Consider an infinite network with nodes distributed in $\Re^2$ following a Poisson distribution with density ρ; and a pair of nodes separated by a Euclidean distance x are directly connected with probability $g(x)$, independent of the event that another distinct pair of nodes are directly connected. Here, $g : \Re^+ \to [0, 1]$ satisfies the conditions of non-increasing monotonicity and integral boundedness [17, pp. 151-152]. As $\rho \to \infty$, a.a.s. the above network in $\Re^2$ has only a unique unbounded component and isolated nodes.

In [6], Philips et al. proved that the average node degree, i.e. the expected number of neighbors of an arbitrary node, must grow logarithmically with the area of the network to ensure that the network is connected, where nodes are placed randomly on a square according to a Poisson point process with a constant density. This result by Philips et al. actually provides a necessary condition on the average node degree required for connectivity. In [5], Xue et al. showed that in a network with a total of n nodes randomly and uniformly distributed on a unit square, if each node is connected to $c \log n$ nearest neighbors with $c \leq 0.074$ then the resulting random network is a.a.s. disconnected as $n \to \infty$; and if each node is connected to $c \log n$ nearest neighbors with $c \geq 5.1774$ then the network is a.a.s. connected as $n \to \infty$. In [8], Balister et al. advanced the results in [5] and improved the lower and upper bounds to $0.3043 \log n$ and $0.5139 \log n$ respectively. In a more recent paper [10] Balister et al. achieved much improved results by showing that there exists a constant c_{crit} such that if each node is connected to $\lfloor c \log n \rfloor$ nearest neighbors with $c < c_{crit}$ then the network is a.a.s. disconnected as $n \to \infty$, and if each node is connected to $\lfloor c \log n \rfloor$ nearest neighbors with $c > c_{crit}$ then the network is a.a.s. connected as $n \to \infty$. In both [8] and [10], the authors considered nodes randomly distributed following a Poisson process of intensity one on a square of area n. In [7], Ravelomanana investigated the critical transmission range for connectivity in 3-dimensional wireless sensor networks and derived similar results as the 2-dimensional results in [1].

In [12], Bettstetter empirically investigated the minimum node degree and connectivity of a finite network with n ($100 \leq n \leq 2000$) nodes randomly and uniformly placed on a square of area A. Tang et al. [13] proposed an empirical formula relating the probability of having a connected network to the transmission range for a finite network with n ($n \leq 125$) nodes randomly and uniformly distributed on a unit square. Bettstetter [11] studied the network connectivity considering different node placement models, i.e. uniform distribution, Gaussian distribution. Note that most results for finite n are empirical results.

B. More general connection models and connectivity

All the work described in the last subsection is based on the unit disk model. This model may simplify analysis but no real antenna has an antenna pattern similar to it. The log-normal shadowing connection model, which is more realistic than the unit disk model, has accordingly been considered for investigating network connectivity in [18]–[23]. Under the log-normal shadowing connection model, two nodes are directly connected if the received power at one node from the other node, whose attenuation follows the log-normal model [24], is greater than or equal to a given threshold.

In [18], Hekmat et al. proposed an empirical formula for computing the average size of the largest connected component through simulations, where a total of n nodes are randomly and uniformly distributed in a bounded area in $\Re^2$. In [22], Bettstetter derived a lower bound on the minimum node density ρ required to ensure that a network with nodes Poissonly distributed in an area in $\Re^2$ with density ρ is k-connected with a high probability. The analysis is based on the observation that the minimum node density required for a k-connected network is larger than that required for the network to have a minimum node degree k, and the assumption that the event that a node has a degree greater than or equal to k is independent of the event that another node has a degree greater than or equal to k. Using simulations, they showed that the bound is tight when the node density is sufficiently large. Using

the same model as in [22], Bettstetter et al. obtained in [23] a lower bound on the minimum node density required for an almost surely connected network using essentially the same technique as that in [22]. The analysis relies on the assumption that the event that a node is isolated and the event that another node is isolated are independent, hereafter referred to as the *independence* assumption. Orriss et al. [19] considered nodes uniformly and randomly distributed on a plane and communicating with each other following the log-normal shadowing model in the framework of cellular networks. They investigated the distribution of the number of base stations that communicate with a given mobile and found that the number of base stations able to communicate with a given mobile *and* lying within a specified range of the mobile follows a Poisson distribution. In [21], Miorandi et al. presented an analytical procedure for computing the node isolation probability in the presence of channel randomness, where nodes are distributed following a Poisson point process in $\Re^2$ (which extends their earlier work in [20]). They further obtained an estimate of the probability that there is no isolated node in the network based on the above independence assumption. The previous results in [18]–[23] dealing with a necessary condition on the critical transmission power for connectivity under the log-normal shadowing model all rely on the independence assumption that the node isolation events are independent, which has only been validated by simulations. Realistically however, one may expect that the event that a node is isolated and the event that another node is isolated will be correlated whenever there is a non-zero probability that a third node may exist which may have direct connections to both nodes. In the unit disk model, this may happen when the transmission range of the two nodes overlaps. In the log-normal model, *any* node may have a non-zero probability of having direct connections to both nodes. This observation and a lack of rigorous analysis on the node isolation events to support the independence assumption raised a question mark over the validity of the results of [18]–[23].

Other work in the area includes [25]–[28], which studies from the percolation perspective, the impact of mutual interference caused by simultaneous transmissions, the impact of physical layer cooperative transmissions, the impact of directional antennas and the impact of unreliable links on connectivity respectively.

C. Random connection model and connectivity

In the more recent work [29], [30], the authors considered a network where all nodes are distributed on a unit square $A \triangleq \left[-\frac{1}{2}, \frac{1}{2}\right]^2$ following a Poisson distribution with known density ρ and a pair of nodes are directly connected following a *random connection model*, viz. a pair of nodes separated by a Euclidean distance x are directly connected with probability $g_{r_\rho}(x) \triangleq g\left(\frac{x}{r_\rho}\right)$, where $g : [0, \infty) \to [0, 1]$, independent of the event that another pair of nodes are directly connected. Here

$$r_\rho = \sqrt{\frac{\log \rho + b}{C\rho}} \tag{1}$$

and b is a constant. The function g is required to satisfy the properties of non-increasing monotonicity and integral boundedness [17], [31, Chapter 6]. Further, it is required that g satisfies the more restrictive requirement that

$$g(x) = o_x\left(\frac{1}{x^2 \log^2 x}\right) \tag{2}$$

in order for the impact of the *truncation effect*, which accounts for the difference between an infinite network and a finite (or asymptotically infinite) network, on connectivity to be asymptotically vanishingly small [29]. Denote the above network by $\mathcal{G}\left(\mathcal{X}_\rho, g_{r_\rho}, A\right)$.

A number of results were obtained based on the connectivity of $\mathcal{G}\left(\mathcal{X}_\rho, g_{r_\rho}, A\right)$:

1) Using the Chen-Stein technique [32], [33], it was shown that as $\rho \to \infty$, the distribution of the number of isolated nodes in $\mathcal{G}\left(\mathcal{X}_\rho, g_{r_\rho}, A\right)$ asymptotically converges to a Poisson distribution with mean e^{-b};
2) The number of isolated nodes due to the boundary effect in $\mathcal{G}\left(\mathcal{X}_\rho, g_{r_\rho}, A\right)$ is a.a.s. zero, i.e. the boundary effect has asymptotically vanishing impact on the number of isolated nodes;
3) As $\rho \to \infty$, the number of components of finite order $k > 1$ in $\mathcal{G}\left(\mathcal{X}_\rho, g_{r_\rho}, A\right)$ asymptotically vanishes;
4) As $\rho \to \infty$, the number of components in $\mathcal{G}\left(\mathcal{X}_\rho, g_{r_\rho}, A\right)$ of unbounded order converges to one;
5) Finally, based on the above results, it was shown that as $\rho \to \infty$, a.a.s. there are only a unique unbounded component and isolated nodes in $\mathcal{G}\left(\mathcal{X}_\rho, g_{r_\rho}, A\right)$, and a sufficient and necessary condition for $\mathcal{G}\left(\mathcal{X}_\rho, g_{r_\rho}, A\right)$ to be a.a.s. connected is that there is no isolated node in the network. Further, the probability that $\mathcal{G}\left(\mathcal{X}_\rho, g_{r_\rho}, A\right)$ has no isolated nodes and the probability that $\mathcal{G}\left(\mathcal{X}_\rho, g_{r_\rho}, A\right)$ forms a connected network both converge to $e^{-e^{-b}}$ as $\rho \to \infty$. As a ready consequence of these results, $\mathcal{G}\left(\mathcal{X}_\rho, g_{r_\rho}, A\right)$ is a.a.s. connected iff $b \to \infty$ as $\rho \to \infty$; and is a.a.s. disconnected iff $b \to -\infty$ as $\rho \to \infty$.

The above results extend the earlier work by Penrose [15], [16] and Gupta and Kumar [1] from the unit disk model to the more generic random connection model and bring theoretical research in the area closer to reality. It can be readily shown that the results on the random connection model include the work of Penrose [15], [16] and Gupta and Kumar [1] on the unit disk model and the work on the log-normal model [18]–[23] as two special cases.

D. Challenges

There remain significant challenges ahead.

Most results in the area, including the results in [29], [30], rely on three main assumptions: a) the connection function g is isotropic, b) the connections are independent, c) nodes are Poissonly or uniformly distributed.

We conjecture that assumption a) is not a critical assumption, i.e. under some mild conditions, e.g. nodes are independently and randomly oriented, assumption a) can be removed while the above results, particularly the ones obtained assuming a random connection model, are still valid. It however remains to validate the conjecture.

The above results however critically rely on assumption b), which is not necessarily valid in some networks due to channel correlation and interference, where the latter effect makes the connection between a pair of nodes dependent on the locations and activities of other nearby nodes. In [34], some preliminary work was conducted on the connectivity of CSMA networks considering the impact of interference. The work essentially uses a decoupling approach to solve the challenges of connection correlation caused by interference by developing an upper bound on the interference experienced by any receiver in the network and then studying the connectivity of the CSMA network using the bound. Their results suggest that when some realistic constraints are considered, i.e. carrier-sensing, the connectivity results, e.g. transmission power required for a connected network, will be very close to those obtained under a unit disk model. This conclusion is in stark contrast with that obtained under an ALOHA multiple-access protocol [25]. Other work in the area includes the work of Haenggi and his colleagues (see e.g. [2], [35]), which characterizes various properties of multi-hop networks subject to interference by using Poisson distribution to approximate the distribution of the set of concurrent transmitters. The major obstacle in dealing with the impact of channel correlation is that there is no widely accepted model in the wireless communication community capturing the impact of channel correlation on connections.

Finally, it is a logical move after the above work to consider connectivity of networks with nodes distributed following a generic distribution other than Poisson or uniform. This remains a major challenge in the area.

III. Connectivity of Giant Component

A *giant component* is a component with a designated large percentage of nodes in the network, say p where $0.5 < p < 1$. A *component* is a maximal set of nodes where there is a path between any pair of nodes in the set.

Results on connectivity of large-scale random networks under both the unit disk model [1], [15], [16] and the more generic random connection model [29], [30] revealed the same scaling law. That is, when the number of nodes, denoted by n, in a network increases, the transmission range (or power) has to increase at a rate to maintain an average node degree of $\Theta(\log n)$ in order to achieve connectivity. For two functions $f(x)$ and $h(x)$, $f(x) = \Theta(h(x))$ iff there exist a sufficiently large x_0 and two positive constants c_1 and c_2 such that for any $x > x_0$, $c_1 h(x) \geq f(x) \geq c_2 h(x)$. For example, the critical transmission range for connectivity is $r(n) = \sqrt{\frac{\log n + c(n)}{\pi n}}$ under the unit disk model for a random network formed by uniformly placing n nodes on a unit-area disk [1], [15], [16][1]. In other words, a connected network poses a very demanding requirement on the transmission range (or power). This in turn causes many undesirable effects on increased interference and reduced throughput.

In the following, we demonstrate the connections between the results on connectivity and the results on network capacity. In [36], it was shown that the end-to-end throughput between a randomly chosen source-destination pair in the above network is $\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$, where W is the link capacity. Refer to [36] for a rigorous definition of network capacity. This result on capacity can be intuitively explained using the results on connectivity as follows: as the number of nodes n increases, the average distance, measured by the number of hops, between a randomly chosen pair of nodes is $\Theta\left(\frac{1}{r(n)}\right) = \Theta\left(\sqrt{\frac{\pi n}{\log n}}\right)$. That is, for a *typical* node, for every packet transmitted for itself, there are $\Theta\left(\sqrt{\frac{\pi n}{\log n}}\right)$ relay packets transmitted for other source-destination pairs. Further, the average node degree is $n\pi r^2(n) = \Theta(\log n)$, which implies that in a neighborhood of a typical node, at any time there can only be one out of every $\Theta(\log n)$ nodes active. It follows that the end-to-end throughput between a typical source-destination pair is $\frac{W}{\Theta\left(\sqrt{\frac{\pi n}{\log n}}\right)\Theta(\log n)} = \Theta\left(\frac{W}{\sqrt{n \log n}}\right)$, hence comes the result in [36]. The above result can be more rigorously derived by following similar steps as those in [36] on analyzing the capacity of a random network. Therefore the reduced capacity as $n \to \infty$ is attributable to the more demanding requirement on the transmission range (or power) to maintain connectivity as $n \to \infty$.

The above observation motivates a question: since the network connectivity is a very demanding requirement, whether there is any benefit in backing down from such a demanding requirement and requiring most nodes, instead all nodes, to be connected?

Indeed in many applications, it is unnecessary for *all* nodes to be *always* connected to each other [37]. Examples of such applications include a wireless sensor network for habitat monitoring [38], [39] or environmental monitoring [40], [41] and a mobile ad-hoc network in which users can tolerate short off-service intervals [42].

In environmental monitoring, there are scenarios where the size of the monitored phenomenon is very large (e.g. rain clouds) or the parameters (e.g. temperature, humidity) that are monitored change slowly both in space and in time. When the number of nodes for monitoring the phenomenon or measuring the parameters is very large,

[1]By scaling, it can be shown that assuming an extended network model where nodes are distributed on a disk of area n with a constant density of 1 node per unit area, the critical transmission range for connectivity is $r(n) = \sqrt{\frac{\log n + c(n)}{\pi}}$.

having a few disconnected nodes will not cause a statistically significant change in the monitored parameters. One example of such applications is a wireless sensor network that was deployed underneath the Briksdalsbreen glacier in Norway to monitor the pressure, humidity, and temperature of ice to understand glacial dynamics in response to climate change [40]. In habitat monitoring, there are scenarios where the number of objects (e.g. zebras and cane toads [38]) that are monitored is large. Having a few nodes disconnected or lost may not significantly affect the accuracy of the monitored parameter. In many mobile ad-hoc networks, having a number of nodes temporarily disconnected is also not critical, as long as users can tolerate short off-service intervals. For example, in a campus-wide wireless network, students and staff can share information using wireless devices (e.g. laptops and personal digital assistants) around the campus [42]. When a wireless device temporarily loses connection, it can store the data and complete the work after becoming connected later.

In [43], [44], considering a network with a total of n nodes uniformly and i.i.d. on a unit square in $\Re^2$, it was shown analytically that under both the unit disk model [44] and the log-normal model [43], the transmission range (or power) required for having a designated large percentage of nodes connected, say p where $0.5 \leq p < 1$, is asymptotically vanishingly small compared to that required for having a connected network, irrespective of the value of p. This result implies that significant energy savings can be achieved if we require only most nodes (e.g. 95%, 99%) to be connected, instead of requiring all nodes to be connected; given a network with most nodes connected, a sharp increase in the transmission range (or power) is required to connect the few remaining hard-to-reach nodes; and the transmission range (or power) required for a large network to be connected is dominated by these hard-to-reach nodes or rare events. It was further shown using simulations that under the unit disk model, in a network with 1000 nodes, the transmission range required for having 95% nodes connected is only 76% of that required for having all nodes connected. Based on a conservative estimate that the required transmission power increases with the square of the required transmission range, an energy saving of at least 42% can be achieved by sacrificing 5% of nodes. That energy saving will further increase with an increase in the number of nodes in the network. Other benefits of the reduced transmission range or power requirement is the reduced interference, hence better throughput.

It remains to find the value of the transmission range (or power) required for guaranteeing a designated large percentage of nodes to be connected in a large scale network. This problem has some intrinsic connections to the problem of finding the percolation probability in the continuum percolation theory [17], which is a well-known open problem in the area. Further, it remains to quantitatively characterize the benefit in capacity due to the reduced transmission range (or power) required for a giant component.

Other researchers approached the problem caused by the demanding requirement of a connected network on the transmission range (or power) from a different perspective and considered the use of infrastructure instead. Here the infrastructure can be quite flexible. It can be a subset of nodes connected through wired connections [45], or a subset of nodes with possibly more powerful transmission capability that forms a wireless backbone of the network [46], [47], or a subset of nodes with satellite links as one would possibly encounter in UAV formations [48]. The use of infrastructure does not change the wireless multi-hop nature of the end-to-end communication, instead the infrastructure assists the end-to-end communication by leapfrogging some long hops and reducing the number of hops between two nodes, hence improving the performance. Accordingly the concept of k-hop connected networks was proposed and investigated [49]–[52]. In a k-hop connected network, the maximum number of hops between any two nodes is smaller than or equal to k. Some research in the area was also conducted under the name of hybrid networks [45], [53].

Despite previous research in the area of hybrid networks or k-hop connected networks, no conclusive results have been obtained yet on the role of infrastructure in wireless multi-hop networks with many problems remain unanswered. Some examples include: for randomly deployed infrastructure nodes and "ordinary" nodes, how many infrastructure nodes (versus ordinary nodes) are required for a k-hop connected network; for deterministically deployed infrastructure nodes and randomly deployed ordinary nodes, how many infrastructure nodes are required for a k-hop connected network and what is the optimum deployment of infrastructure nodes; how to combine the use of infrastructure-based communications and ad-hoc communications in one network in order to provide some performance guarantee, in terms of capacity or delay. These problems are important for wireless multi-hop networks to provide reliable services, particularly for wireless vehicular networks in which both infrastructure-based communications and ad-hoc communications will co-exist [54].

IV. Development and Challenges in Mobile Networks

In [55], Grossglauser and Tse studied the capacity of mobile ad-hoc networks. Particularly, they considered a network with a total of n nodes distributed on a unit-area disk, the trajectories of different nodes are i.i.d. and the nodes' movement is such that the spatial distribution of nodes are stationary and ergodic with stationary uniform distribution on the disk. They showed that in the above network with *unbounded delay requirement*, the throughput between a randomly chosen source-destination pair can be kept *constant* even as n increases. This result is in stark contrast with its counter-part in static networks in which the throughout between a randomly chosen source-destination pair is shown to be $\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ [36].

Following the seminal work of Grossglauser and Tse, other researchers have conducted further research trying to quantitatively characterize the relationship between delay, mobility and capacity in mobile ad-hoc networks [49], [56]–[59] and the obtained results vary greatly with the mobility models and network settings.

A fundamental reason why mobility increases throughput is that in mobile networks message transmissions generally follow the store-carry-forward pattern versus the store-forward pattern found in static networks. As nodes move, new opportunity may arise such that a mobile node can carry the message until it meets a node, which is in a better position than itself to transmit the message to the destination, or until it meets the destination directly. In this way, the number of relay nodes (number of hops) involved in transmitting a message to its destination can be greatly reduced and the required transmission range (or power) for a node to reach another node via a multi-hop path can also be greatly reduced, hence the benefit in improved capacity. The cost in achieving this benefit in capacity is the increased delay.

Following the techniques demonstrated in Section III, the results in [55] on the capacity of mobile networks can also be obtained as follows. In the network model considered in [55], a two-hop relaying strategy is employed. Therefore as the number of nodes n increases, the average distance, measured by the number of relay hops, between a randomly chosen pair of source-destination is bounded by 2. Further, the value of the transmission range r does not have to increase with n because connectivity (in the sense that every pair of nodes can exchange packets) can be achieved through node movement. Therefore the average node degree is also $\Theta(1)$. It then readily follows that the capacity of the mobile network is $\Theta(1)$. The end-to-end delay can be analyzed by evaluating the time that two balls with radius equal to $r/2$, representing the source and the relay node respectively, hit a (randomly chosen) third ball, representing the destination, with radius $r/2$.

The above observation motivates us to conclude that the fundamental factors that determine the capacity of a mobile (or static) network are:

1) The expected number of simultaneously active transmissions. This is further determined by the spatial node distribution and the transmission range (or power).
2) The average number of relay hops between a source and its destination. It determines the average number of times that a packet need to be transmitted before reaching its destination, i.e. the transmission capacity consumed for an end-to-end transmission.

Therefore the only difference between mobile and static networks is that some part of the job involved in moving a packet, originally taken care of entirely by wireless transmissions in static network, can now be taken care of by the physical movement of nodes in mobile networks. By viewing both the physical movements of nodes and the wireless transmission as simply a way to move packets physically over a distance, a unified theory for analyzing the performance of both mobile and static networks can be established.

By analogy, mobility can also improve connectivity. There are three fundamental differences between mobile networks and static networks [60] from a graph theory perspective: in mobile networks

- the wireless link between two directly connected nodes and the end-to-end path only exists temporarily;
- two nodes may never be part of the same connected component but they are still able to communicate, i.e. exchange messages, with each other; and
- while any one wireless link may be (or assumed to be) undirectional, the path connecting any two nodes is directional, i.e. there is a path from node v_i to node v_j within a designated time period does not necessarily mean there is a path from v_j to v_i within the same period.

These are illustrated in Fig. 1. Particularly the last difference implies that it is important to consider the temporal order of links when analyzing mobile networks, which has been incorrectly neglected in some previous work.

Due to these differences, many established concepts in static networks must be revisited for mobile networks. For example, a static wireless multi-hop network is said to be connected iff there is a path between any pair of nodes in the network. However a more meaningful definition of connectivity in mobile networks is to say that a mobile network is connected in time period $[0, T]$ if any node can exchange a message with any other node within $[0, T]$. The above definition implies that the tradeoff between connectivity, mobility and delay is the prime issue when analyzing the connectivity of mobile networks. Despite intensive research on the properties of mobile networks, no conclusive results have been obtained on the above problem and it remains a major challenge in the area.

V. Summary

Wireless multi-hop networks have attracted significant research interest. This interest is expected to grow further with the proliferation of applications, particularly in the areas of wireless vehicular networks and sensor networks. In this paper, we briefly overviewed recent development and discussed research challenges and opportunities in the area mainly from the perspective of network connectivity. We also showed how the results on network connectivity are related to the study of other performance metrics, i.e. capacity and delay. We pointed out the fundamental parameters determining the capacity of a wireless multi-hop network and the fundamental difference between mobile and static networks.

References

[1] P. Gupta and P. R. Kumar, *Critical Power for Asymptotic Connectivity in Wireless Networks*. Boston, MA: Birkhauser, 1998, pp. 547–566.

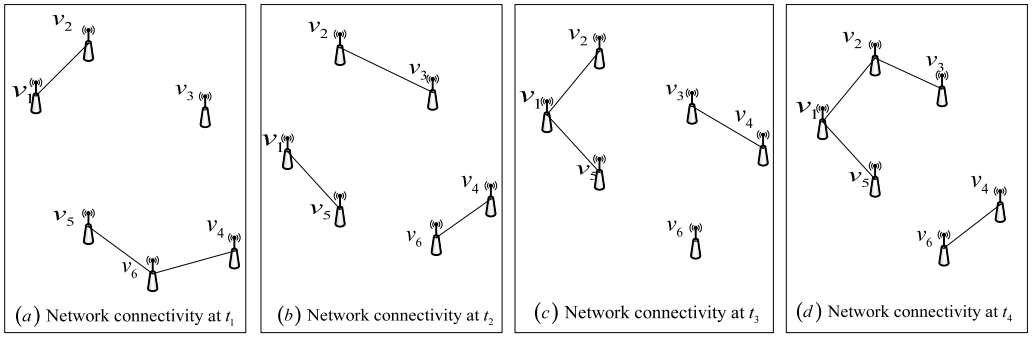


Figure 1. An illustration of connectivity in a mobile ad-hoc network. A solid line represents a connection between two nodes. The network is disconnected at any time instant but there is a path from any node to any other node in the network. For example, nodes v_1 and v_6 are never part of the same connected component but a message from v_1 can still reach v_6 through the following path: $t_1 : v_1 \rightarrow v_2$, $t_2 : v_2 \rightarrow v_3$, $t_3 : v_3 \rightarrow v_4$, $t_4 : v_4 \rightarrow v_6$. Further, a message from v_6 can reach v_1 at t_2 but a message from v_1 can only reach v_6 at t_4.

[2] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, 2009.

[3] M. D. Penrose, *Random Geometric Graphs*, ser. Oxford Studies in Probability. Oxford University Press, USA, 2003.

[4] ——, "On k-connectivity for a geometric random graph," *Random Structures and Algorithms*, vol. 15, no. 2, pp. 145–164, 1999.

[5] F. Xue and P. Kumar, "The number of neighbors needed for connectivity of wireless networks," *Wireless Networks*, vol. 10, no. 2, pp. 169–181, 2004.

[6] T. K. Philips, S. S. Panwar, and A. N. Tantawi, "Connectivity properties of a packet radio network model," *IEEE Transactions on Information Theory*, vol. 35, no. 5, pp. 1044–1047, 1989.

[7] V. Ravelomanana, "Extremal properties of three-dimensional sensor networks with applications," *IEEE Transactions on Mobile Computing*, vol. 3, no. 3, pp. 246–257, 2004.

[8] P. Balister, B. Bollobas, A. Sarkar, and M. Walters, "Connectivity of random k-nearest-neighbour graphs," *Advances in Applied Probability*, vol. 37, no. 1, pp. 1–24, 2005.

[9] P.-J. Wan and C.-W. Yi, "Asymptotic critical transmission radius and critical neighbor number for k-connectivity in wireless ad hoc networks," 2004.

[10] P. Balister, B. Bollobas, A. Sarkar, and M. Walters, "A critical constant for the k nearest neighbour model," *Advances in Applied Probability*, vol. 41, no. 1, pp. 1–12, 2009.

[11] C. Bettstetter, "On the connectivity of ad hoc networks," *The Computer Journal*, vol. 47, no. 4, pp. 432–447, 2004.

[12] ——, "On the minimum node degree and connectivity of a wireless multihop network," in *3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pp. 80–91.

[13] A. Tang, C. Florens, and S. H. Low, "An empirical study on the connectivity of ad hoc networks," in *IEEE Aerospace Conference*, vol. 3, pp. 1333–1338.

[14] C. Bettstetter and J. Zangl, "How to achieve a connected ad hoc network with homogeneous range assignment: an analytical study with consideration of border effects," in *4th International Workshop on Mobile and Wireless Communications Network*, pp. 125–129.

[15] M. Penrose, "The longest edge of the random minimal spanning tree," *The Annals of Applied Probability*, vol. 7, no. 2, pp. 340–361, 1997.

[16] ——, "A strong law for the longest edge of the minimal spanning tree," *The Annals of Applied Probability*, vol. 27, no. 1, pp. 246–260, 1999.

[17] R. Meester and R. Roy, *Continuum Percolation*, ser. Cambridge Tracts in Mathematics. Cambridge University Press, 1996.

[18] R. Hekmat and P. V. Mieghem, "Connectivity in wireless ad-hoc networks with a log-normal radio model," *Mobile Networks and Applications*, vol. 11, no. 3, pp. 351–360, 2006.

[19] J. Orriss and S. K. Barton, "Probability distributions for the number of radio transceivers which can communicate with one another," *EEE Transactions on Communications*, vol. 51, no. 4, pp. 676–681, 2003.

[20] D. Miorandi and E. Altman, "Coverage and connectivity of ad hoc networks presence of channel randomness," in *IEEE INFOCOM*, vol. 1, pp. 491–502.

[21] D. Miorandi, "The impact of channel randomness on coverage and connectivity of ad hoc and sensor networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 1062–1072, 2008.

[22] C. Bettstetter, "Failure-resilient ad hoc and sensor networks in a shadow fading environment," in *IEEE/IFIP International Conference on Dependable Systems and Networks*.

[23] C. Bettstetter and C. Hartmann, "Connectivity of wireless multihop networks in a shadow fading environment," *Wireless Networks*, vol. 11, no. 5, pp. 571–579, 2005.

[24] T. S. Rappaport, *Wireless Communications: Principles and Practice*, ser. Prentice Hall Communications Engineering and Emerging Technologies Series. Prentice Hall, 2002.

[25] O. Dousse, F. Baccelli, and P. Thiran, "Impact of interferences on connectivity in ad hoc networks," *IEEE/ACM Transactions on Networking*, vol. 13, no. 2, pp. 425–436, 2005.

[26] D. Goeckel, L. Benyuan, D. Towsley, W. Liaoruo, and C. Westphal, "Asymptotic connectivity properties of cooperative wireless ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1226–1237, 2009.

[27] P. Li, C. Zhang, and Y. Fang, "Asymptotic connectivity in wireless ad hoc networks using directional antennas," *IEEE/ACM Transactions on Networking*, vol. 17, no. 4, pp. 1106–1117, 2009.

[28] Z. Kong and E. M. Yeh, "Connectivity and latency in large-scale wireless networks with unreliable links," in *IEEE INFOCOM*, pp. 11–15.

[29] G. Mao and B. D. Anderson, "Towards a better understanding of large scale network models," *accepted to appear in IEEE/ACM Transactions on Networking*, 2011.

[30] G. Mao and B. Anderson, "Connectivity of large wireless

networks under a generic connection model," *submitted to IEEE Transactions on Information Theory*, 2012.

[31] M. Franceschetti and R. Meester, *Random Networks for Communication*. Cambridge University Press, 2007.

[32] R. Arratia, L. Goldstein, and L. Gordon, "Poisson approximation and the chen-stein method," *Statistical Science*, vol. 5, no. 4, pp. 403–434, 1990.

[33] A. D. Barbour, L. Holst, and S. Jason, *Poisson Approximation*. Oxford University Press, New York, 2003.

[34] T. Yang, G. Mao, and W. Zhang, "Connectivity of wireless csma multi-hop networks," in *IEEE International Conference on Communications*, pp. 1–5.

[35] J. G. Andrews, S. Weber, M. Kountouris, and M. Haenggi, "Random access transport capacity," *IEEE Transactions on Wireless Communications*, vol. 9, no. 6, pp. 2101 – 2111, 2010.

[36] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[37] D. M. Blough, M. Leoncini, G. Resta, and P. Santi, "The k-neighbors approach to interference bounded and symmetric topology control in ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 9, pp. 1267 – 1282, 2006.

[38] W. Hu, V. N. Tran, N. Bulusu, C. Chou, S. Jha, and A. Taylor, "The design and evaluation of a hybrid sensor network for cane-toad monitoring," in *Proc. 4th Int. Symp. Inf. Process. Sens. Netw.*, pp. 503 – 508.

[39] P. Juang, H. Oki, Y. Wang, M. Martonosi, L. Peh, and D. Rubenstein, "Energy-efficient computing for wildlife tracking: Design tradeoffs and early experiences with zebranet," *ACM SIGARCH Comput. Archit. News*, vol. 30, no. 5, pp. 96 – 107, 2002.

[40] P. Padhy, K. Martinez, A. Riddoch, J. K. Hart, and H. L. R. Ong, "Glacial environment monitoring using sensor networks," in *Proc. 1st REALWSN*, pp. 10 – 14.

[41] D. Ingraham, R. Beresford, K. Kaluri, M. Ndoh, and K. Srinivasan, "Wireless sensors: Oyster habitat monitoring in the bras d'or lakes," in *IEEE 1st International Conference on Distributed Computing In Sensor Systems*, pp. 399 – 400.

[42] X. Yu and S. Chandra, "Delay-tolerant collaborations among campus wide wireless users," in *IEEE INFOCOM*, pp. 2101 – 2109.

[43] X. Ta, G. Mao, and B. D. Anderson, "On the giant component of wireless multi-hop networks in the presence of shadowing," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 9, pp. 5152–5163, 2009.

[44] ——, "On the properties of giant component in wireless multi-hop networks," in *IEEE INFOCOM*, pp. 2556 – 2560.

[45] B. Liu, Z. Liu, and D. Towsley, "On the capacity of hybrid wireless networks," in *IEEE INFOCOM*, vol. 2, pp. 1543–1552.

[46] M. Franceschetti, O. Dousse, D. N. C. Tse, and P. Thiran, "Closing the gap in the capacity of wireless networks via percolation theory," *IEEE Transactions on Information Theory*, vol. 53, no. 3, pp. 1009–1018, 2007.

[47] P. Li and Y. Fang, "The capacity of heterogeneous wireless networks," in *IEEE INFOCOM*, pp. 1–9.

[48] G. Mao, S. Drake, and B. D. O. Anderson, "Design of an extended kalman filter for uav localization," in *Information, Decision and Control*, pp. 224–229.

[49] Q. Wang, X. Wang, and X. Lin, "Mobility increases the connectivity of k-hop clustered wireless networks," pp. 121 – 132, 2009.

[50] G. Mao, Z. Zhang, and B. Anderson, "Probability of k-hop connection under random connection model," *IEEE Communication Letters*, vol. 14, no. 11, pp. 1023 – 1025, 2010.

[51] S. C. Ng, W. Zhang, Y. Yang, and G. Mao, "Analysis of access and connectivity probabilities in vehicular relay networks," *IEEE Journal on Selected Areas in Communications–Special Issue Vehicular Communications and Networks*, vol. 29, no. 1, pp. 140 – 150, 2011.

[52] X. Ta, G. Mao, and B. D. O. Anderson, "Evaluation of the probability of k-hop connection in homogeneous wireless sensor networks," pp. 1279 – 1284, 2007.

[53] O. Dousse, P. Thiran, and M. Hasler, "Connectivity in ad-hoc and hybrid networks," in *IEEE INFOCOM*, vol. 2, pp. 1079–1088.

[54] Y. Yang, H. Hu, J. Xu, and G. Mao, "Relay technologies for wimax and lte-advanced mobile systems," *IEEE Communications Magazine*, vol. 47, no. 10, pp. 100–105, 2009.

[55] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 477–486, 2002.

[56] G. Sharma, R. Mazumdar, and B. Shroff, "Delay and capacity trade-offs in mobile ad hoc networks: A global perspective," *IEEE/ACM Transactions on Networking*, vol. 15, no. 5, pp. 981–992, 2007.

[57] M. J. Neely and E. Modiano, "Capacity and delay tradeoffs for ad hoc mobile networks," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1917–1937, 2005.

[58] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-delay trade-off in wireless networks," in *IEEE INFOCOM*, vol. 1, p. 475.

[59] S. Toumpis and A. J. Goldsmith, "Large wireless networks under fading, mobility, and delay constraints," in *INFOCOM*, vol. 1, pp. 609–619.

[60] G. Mao and B. D. Anderson, "Graph theoretic models and tools for the analysis of dynamic wireless multihop networks," in *IEEE WCNC*, pp. 1–6.

Analysis of Access and Connectivity Probabilities in Vehicular Relay Networks

Seh Chun Ng, *Student Member, IEEE,* Wuxiong Zhang, Yu Zhang, Yang Yang, *Senior Member, IEEE,* and Guoqiang Mao, *Senior Member, IEEE*

***Abstract*—IEEE 802.11p and 1609 standards are currently under development to support Vehicle-to-Vehicle and Vehicle-to-Infrastructure communications in vehicular networks. For infrastructure-based vehicular relay networks, access probability is an important measure which indicates how well an arbitrary vehicle can access the infrastructure, i.e. a base station (BS). On the other hand, connectivity probability, i.e. the probability that all the vehicles are connected to the infrastructure, indicates the service coverage performance of a vehicular relay network. In this paper, we develop an analytical model with a generic radio channel model to fully characterize the access probability and connectivity probability performance in a vehicular relay network considering both one-hop (direct access) and two-hop (via a relay) communications between a vehicle and the infrastructure. Specifically, we derive close-form equations for calculating these two probabilities. Our analytical results, validated by simulations, reveal the tradeoffs between key system parameters, such as inter-BS distance, vehicle density, transmission ranges of a BS and a vehicle, and their collective impact on access probability and connectivity probability under different communication channel models. These results and new knowledge about vehicular relay networks will enable network designers and operators to effectively improve network planning, deployment and resource management.**

***Index Terms*—Vehicular Ad Hoc Network (VANET), Wireless Access in Vehicular Environments (WAVE), IEEE 802.11p, IEEE 1609, access probability, connectivity, relay.**

Manuscript received 5 January 2010; revised 7 May 2010 and 12 July 2010. This work is partially supported by the Australian Research Council (ARC) under the Discovery project DP0877562, the National Natural Science Foundation of China (NSFC) under the grant 60902041, the Ministry of Science and Technology (MOST) of China under the grant 2009DFB13080, and the RCUK project "UK-China Science Bridges: R&D on (B)4G Wireless Mobile Communications (EP/G042713/1)". This work is partially supported by National ICT Australia (NICTA), which is funded by the Australian Government Department of Communications, Information Technology and the Arts and the Australian Research Council through the Backing Australia Ability initiative and the ICT Centre of Excellence Program, and by the Air Force Research Laboratory, under agreement number FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

S. C. Ng and G. Mao are with School of Electrical and Information Engineering, University of Sydney, Australia and National ICT Australia (NICTA), Sydney (e-mail: {sehchun,guoqiang}@ee.usyd.edu.au).

W. Zhang is with Shanghai Research Center for Wireless Communications (WiCO), Graduate School of the Chinese Academy of Sciences, China (e-mail: wuxiong.zhang@shrcwc.org).

Y. Zhang is with the Dept. of Computing, Imperial College London, UK (e-mail: yuzhang@doc.ic.ac.uk).

Y. Yang is with Shanghai Research Center for Wireless Communications (WiCO), Shanghai Institute of Microsystem and Information Technology (SIMIT), Chinese Academy of Sciences, China (e-mail: yang.yang@shrcwc.org) corresponding author.

Digital Object Identifier 10.1109/JSAC.2011.110114.

I. Introduction

VEHICULAR ad-hoc network (VANET) is a type of promising application-oriented network deployed along a highway for safety and emergency information delivery (for drivers), entertainment content distribution (for passengers), and data collection and communication (for road and traffic managers). VANET is a hybrid wireless network that supports both infrastructure-based and ad hoc communications. Specifically, vehicles on the road can communicate with each other through a multi-hop ad hoc connection. They can also access the Internet and other broadband services through the roadside infrastructure, i.e. base stations (BSs) or access points (APs) along the road. When a vehicle moves out of the radio coverage area of a BS, e.g. it is located in the coverage gap between two adjacent BSs, it will identify and use its neighboring vehicles (if any) as relays to access the roadside infrastructure. These types of Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) communications have recently received significant interests from both academia and industry [1], [2], [3]. V2V communication has so far been envisioned for supporting safety and traffic management applications. With better sensing and data communication techniques, drivers can share the information such as slippery road, poor visibility, sudden stop and road congestion with each other. Hence warnings are provided to prevent accidents and improve road safety.

As shown in Fig. 1, IEEE 802.11p standard cooperates with the IEEE 1609 standard family, which is developed to support Wireless Access in Vehicular Environment (WAVE) and to deliver safety and infotainment applications to vehicles on the road [4], [5]. Specifically, IEEE 802.11p is a draft amendment (with WAVE capability) to the IEEE 802.11 standards, which is expected to be finalized and approved in 2010. The goal of 802.11p standard is to provide V2V and V2I communications, up to a range of 1km, at an average data rate of 6 Mbps over the dedicated 5.9 GHz (5.85-5.925 GHz) licensed frequency band. IEEE 802.11p uses an amended 802.11a physical-layer specification with Orthogonal Frequency-Division Multiplexing (OFDM) technique. On Medium Access Control (MAC) layer, it adopts the Enhanced Distributed Channel Access (EDCA) protocol from the 802.11e standard to support Quality of Service (QoS). The following standards are included in the IEEE 1609 standard family: IEEE P1609.0, IEEE P1609.1, IEEE P1609.2, IEEE P1609.3, IEEE P1609.4. Some new standards have recently been added to IEEE 1609 family, such as 1609.5 (Communications Management), 1609.6 (Facilities) and 1609.11 (Electronic Payment Service). Their functions and relationships with other 1609 standards are shown in Fig. 1.

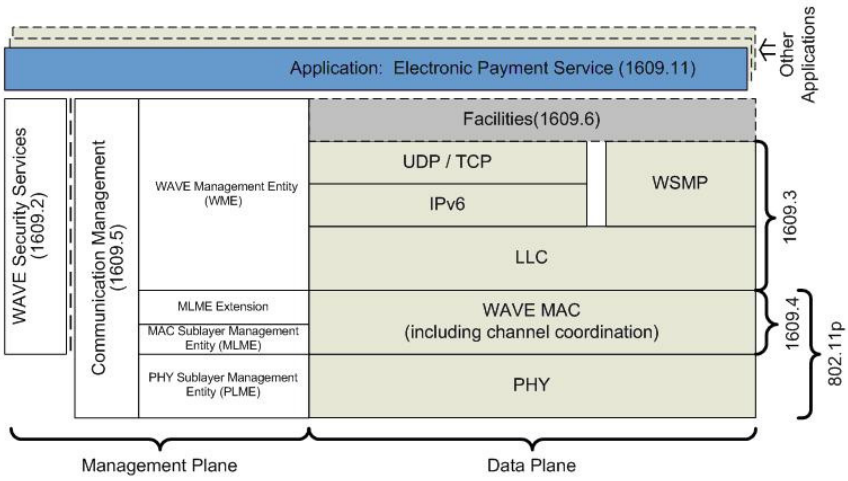


Fig. 1. Architecture of IEEE 1609 standards family

IEEE 802.11p and 1609 standards are still in the draft stage. The harsh vehicular communication environment, caused by variable vehicle speed, high mobility, and dynamic network topology, brings many technical challenges in developing and deploying WAVE applications and services. From a user/vehicle's perspective, the first, and probably the most important, service requirement is to be able to access the roadside infrastructure, directly or indirectly via a relay vehicle. From the perspective of a network operator or service provider, it is important to guarantee satisfactory and profitable service coverage while minimizing the deployment and maintenance costs of the roadside infrastructure.

To improve user satisfaction and service coverage of future IEEE 1609 based WAVE systems and applications, this research develops an analytical model to fully characterize the access probability (for user satisfaction analysis) and the connectivity probability (for service coverage analysis) for infrastructure-based vehicular relay networks, wherein both one-hop (direct access) and two-hop (via a relay) communications between a vehicle and the infrastructure (i.e. a BS) are supported. In this paper, a generic connection model is used to investigate the impact of different system parameters, i.e., inter-BS distance (or BS density), vehicle density, radio coverage ranges of BSs and vehicles, on key performance metrics, i.e. user access probability and service connectivity probability. The analysis is then applied to two widely used communication channel models as specific examples of the generic connection model. This research enables us to improve access probability and connectivity probability in vehicular relay networks, and therefore support reliable V2I and V2V data transmissions in different commercial applications and services, such as emergency messaging service, mobile Internet access and on-road entertainments.

The rest of this paper is organized as follows. In Section II we introduce related work. In Section III we define the system model. In Section IV we present the analysis of the access and connectivity probabilities under a generic radio channel model. In Section V we focus on two widely used radio channel models, i.e. the unit disk communication model and the log-normal shadowing model, and their analysis as special cases of the generic channel model. In Section VI we discuss the analytical and simulation results, followed by conclusions in Section VII.

II. Related Work

Recently, significant research on VANET, WAVE, IEEE 802.11p and 1609 standards has been undertaken to measure, estimate and characterize wireless vehicular channels [6], [7], [8], to model and analyze system performance [9], [10], [11], [12], [13], [14], to design and evaluate MAC protocols [15], [16], [17], [18], [19], and to develop VANET simulator [20] and testbed [21]. Specifically, in [9], it is found a communication distance of 1000m, which is specified in the IEEE802.11p Project Authorization Request (PAR), cannot be achieved by an Equivalent Isotropically Radiated Power (EIRP) of 2W in a vehicle-to-vehicle (V2V) highway scenario. The impacts of vehicle density (or V2V distance) and Line-Of-Sight (LOS) communication link on different system performance metrics, such as throughput, average delay, packet loss, and collision probability, are investigated in [10], [11], [12].

In [15], it is pointed out that the performance of IEEE 802.11p standard is not satisfactory in the infrastructure data collection mode with a static backoff scheme. The capability of 802.11p MAC protocol is further evaluated and enhanced to support both safety applications (i.e. emergency message dissemination with strict time constraints) and non-safety applications [16], [17], [22], [18], [19]. In [23], Salhi et al. presented a novel data gathering and dissemination architecture based on hierarchical and geographical mechanisms for vehicular sensor networks. In [24] and [25] practical traffic prioritization and power control schemes are proposed and evaluated respectively, to support real-time delivery of safety-critical emergency information. In [26], Shrestha et al. develop a new scheme using the BitTorrent tool and bargaining game to efficiently distribute a large amount of data over V2V and V2I communication links.

Access and connectivity probabilities have been studied in the literature for one-dimensional (1-D) [27], [28] and two-dimensional (2-D) [29], [30], [31] multi-hop wireless networks. In [27], Wu focuses on V2V communications and derives a close-form expression of connectivity probability in a linear VANET with high-speed vehicles and time-varying

vehicle populations, i.e. dynamic network topology and vehicle density. The impacts of some key network parameters, such as vehicle arrival rate, random vehicle speed and transmission range, are considered in his work. Based on a Poisson assumption of node distribution, Dousse, Thiran and Hasler study a 1-D network with equally-spaced BSs and Poissonly distributed vehicles in [28]. Considering a unit disk communication model, they derive the connectivity probability (defined as access probability in this paper) that an arbitrary vehicle can reach at least one BS over multiple hops.

For a 2-D multi-hop cellular network, where nodes are uniformly distributed in a circular area of unit radius, Ojha et al. obtain the minimum transmission range required for these nodes to be able to access a BS located at the center of this circular area over multiple hops under the unit disk communication model as the number of nodes goes to infinity [29]. When both BSs and nodes are Poissonly distributed in a 2-D area and a log-normal shadowing communication model is considered, a lower bound on the probability that a node cannot access any BS within a designated number of hops has been derived in [30]. The result relies on the independence assumption that the event that a node cannot access to a BS in a specific number of hops, say t, and the event that another node cannot access to a BS in t hops are independent, which is not always valid for some cases (to be shown in this paper). In [31], the probability that a wireless ad-hoc network with randomly and uniformly distributed nodes form a connected network is studied. It is shown that the probability of having a connected network and the probability of having no isolated node asymptotically converges to the same value as the number of nodes in the network goes to infinity.

Different from previous work carried out mainly under the unit disk communication model and considering no restriction on the maximum hop count for packet transmission, this research focuses on vehicular relay networks with a maximum hop count of two hops to ensure high communication quality and reliability between a node and a BS, i.e. two-hop V2I communications, which is more practical for real-world VANET application scenarios. In addition, our analytical approach uses a generic communication channel model and derives the exact close-form equations of access and connectivity probabilities, not the asymptotic results that are valid only when the number of nodes (or node density) in a network is very large. The results obtained under the generic channel model are then applied to two widely used models, i.e. the unit disk model and the log-normal model, as special cases. Finally, we investigate the impacts of some key system parameters, such as inter-BS distance (or BS density), vehicle density, transmission ranges of a BS and a vehicle, on user access probability and service connectivity probability performance under different communication channel models.

III. System Model

We consider an infrastructure-based vehicular relay network, as shown in Fig. 2, wherein a number of BSs are uniformly deployed along a long road, while other vehicles are distributed on the road randomly according to a Poisson distribution. We analyze the access probability, i.e. the probability that an arbitrary vehicle can access its nearby BSs within two hops, and the connectivity probability, i.e. the probability that all vehicles can access at least one BS within two hops, of the network by investigating a subnetwork bounded by two adjacent base stations. Let L be the Euclidean distance (in meters) between two adjacent BSs and ρ be the vehicle density measured in vehicles per meter (vpm). Since the vehicles are assumed to be Poissonly distributed with density ρ, discussion on the distribution of the number of vehicles on the road is only meaningful if we restrict to the (random) number of vehicles in a specific section of the road, and we call any section of the road a *road segment*. For a road segment with length x, the number of vehicles on that road segment is then a Poisson random variable with mean ρx. So the probability that there are k vehicles on a road segment of x meters is given as

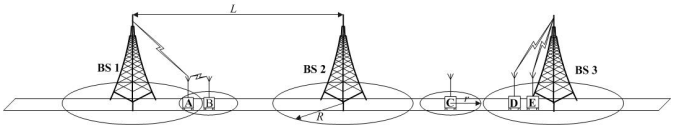


Fig. 2. An Infrastructure-based Vehicular Relay Network.

$$f(k,x) = \frac{(\rho x)^k e^{-\rho x}}{k!}, k \geq 0. \quad (1)$$

Since we investigate a subnetwork bounded by two adjacent BSs, the probability that there are k vehicles on the road segment bounded by two adjacent BSs is then $f(k, L)$.

Assuming a generic channel model $\mathcal{C}$, let $g_v^{\mathcal{C}}(x)$ be the probability that two vehicles separated by an Euclidean distance x are directly connected. Similarly, denote by $g_b^{\mathcal{C}}(x)$ the probability that a vehicle and a BS separated by an Euclidean distance x are directly connected. We assume that the event that two vehicles (or a vehicle and a BS) are directly connected is independent of the event that another two vehicles (or a vehicle and a BS) are directly connected. That is, the event that two vehicles (or in the similar case, between a vehicle and a BS) are directly connected is only determined by the locations of the two vehicles and is not affected by the presence or absence of connections between other pairs of vehicles[1]. We also assume that $g_b^{\mathcal{C}}(x) \geq g_v^{\mathcal{C}}(x)$. This assumption is justified because it is often the case that a BS can not only transmit at a larger transmission power than a vehicle, it can also be equipped with more sophisticated antennas, which make it more sensitive to the transmitted signal from a vehicle.

IV. Analysis of Access and Connectivity Probabilities

Assume that the subnetwork being considered is placed at $[0, L]$. The two BSs at both ends of the subnetwork are labeled

[1]Although field measurements in real applications seem to indicate that the connectivity between different pairs of geographically / frequency proximate wireless nodes are correlated [32], [33], [34], the independence assumption is generally considered appropriate for far-field transmission and has been widely used in the literature under many channel models including log-normal shadowing model [35], [36], [30], [37]. Note that the unit disk model is a special channel model which fulfills the independence assumption by nature [38, pg. 12]. This is because for the unit disk model, two vehicles are directly connected if and only if their Euclidean distance is smaller than the transmission range, and is not affected by the presence or absence of other connections (vehicles).

as BS1 and BS2 and are at 0 and L respectively. Denote by $G(L,\rho,\mathcal{C})$ the subnetwork with length L, vehicle density ρ and channel model $\mathcal{C}$. We investigate the access probability p_a that an arbitrary vehicle in $G(L,\rho,\mathcal{C})$ can access either BS (either BS1 or BS2). We also investigate the probability p_c that all vehicles in $G(L,\rho,\mathcal{C})$ are connected to at least one of the BSs at both ends of the subnetwork.

A vehicle is said to be located at x if its Euclidean distance to BS1 is x. The probability that a vehicle located at x is not directly connected to BS1 and BS2 are $1-g_b^{\mathcal{C}}(x)$ and $1-g_b^{\mathcal{C}}(L-x)$ respectively. Because the event that a vehicle is not directly connected to BS1 and the event that the same vehicle is not directly connected to BS2 are independent, the probability that the vehicle is directly connected to either BS1 or BS2 is then

$$p_1(x) = 1-(1-g_b^{\mathcal{C}}(x))(1-g_b^{\mathcal{C}}(L-x)). \tag{2}$$

In order to derive p_a and p_c we need the following lemmas.

Lemma 1. *Let K_1 be the set of vehicles in the subnetwork $G(L,\rho,\mathcal{C})$ which are directly connected to either BS1 or BS2, then K_1 has an inhomogeneous Poisson distribution with density $\rho p_1(x)$ where $p_1(x)$ is given by Eq.* (2).

Proof: Let K denotes the set of vehicles in $G(L,\rho,\mathcal{C})$. Then K has a homogeneous Poisson distribution with density ρ over the segment $[0,L]$. Consider a realization of K and remove a vehicle located at x from this realization with probability $1-p_1(x)$, independent of the removal probability of other vehicles. The remaining set of vehicles can be effectively viewed as a realization of K_1. Note that the above procedure which removes/retains vehicles independently with some probabilities is called *thinning* [38]. Let $\mathcal{N}(K_1)$ be the number of vehicles in K_1. Then following the law of total probability

$$\begin{aligned}&\Pr(\mathcal{N}(K_1)=j)\\ &=\sum_{i=j}^{\infty}\Pr(\mathcal{N}(K)=i)\Pr(\mathcal{N}(K_1)=j|\mathcal{N}(K)=i).\end{aligned} \tag{3}$$

For a randomly chosen vehicle in K, the vehicle is known to be uniformly distributed in $[0,L]$. Hence, the probability that a vehicle is in K_1 given that the vehicle is in K is

$$q=\frac{1}{L}\int_0^L p_1(x)dx. \tag{4}$$

Since the probability of a randomly chosen vehicle in K being directly connected to either BS1 or BS2 are identically and independently distributed, the probability that among i vehicles in K there are j vehicles in K_1 follows the binomial distributed $B(i,q)$. Thus

$$\Pr(\mathcal{N}(K_1)=j|\mathcal{N}(K)=i)=\binom{i}{j}q^j(1-q)^{i-j}. \tag{5}$$

Applying Eq. (1) and (5) into Eq. (3) we have

$$\begin{aligned}\Pr(\mathcal{N}(K_1)=j)&=\sum_{i=j}^{\infty}\frac{(\rho L)^i}{i!}e^{-\rho L}\binom{i}{j}q^j(1-q)^{i-j}\\ &=\frac{(\int_0^L\rho p_1(x)dx)^j}{j!}e^{-\int_0^L\rho p_1(x)dx}.\end{aligned} \tag{6}$$

Furthermore, denote by $\mathcal{N}(K_1(l))$ the number of vehicles in a road segment l within $[0,L]$ which are directly connected to at least one BSs. Using the above procedure, it is trivial to show that

$$\Pr(\mathcal{N}(K_1(l))=j)=\frac{(\int_{x\in l}\rho p_1(x)dx)^j}{j!}e^{-\int_{x\in l}\rho p_1(x)dx}. \tag{7}$$

For n mutually disjoint road segments $l_1,l_2,\cdots,l_n$ in $[0,L]$, the random variables $\mathcal{N}(K_1(l_1)),\cdots,\mathcal{N}(K_1(l_n))$ are mutually independent. This is because the event that one vehicle is directly connected to either BS1 or BS2 is not affected by the locations of other vehicles, and whether or not those vehicles are directly connected to either BS1 or BS2. Consequently, the existence and locations of vehicles in one road segment will not affect the number of vehicles directly connected to BS1 or BS2 in another disjoint road segment. With the above independence property, Eq. (6) and (7), the proof is then complete. Note that some parts of the proof are similar to the arguments used in [38]. ■

Lemma 2. *Let $p_2(x)$ be the probability that a vehicle located at x in $G(L,\rho,\mathcal{C})$ is directly connected to at least one vehicle in K_1, then*

$$p_2(x)=1-e^{-\int_0^L g_v^{\mathcal{C}}(\|x-y\|)\rho p_1(y)dy} \tag{8}$$

where $p_1(y)$ is given by Eq. (2) *and $\|.\|$ denotes the Euclidean norm.*

Proof: Imagine we partition $[0,L]$ into L/dy non-overlapping intervals of differential length dy. Since dy is a very small value, the probability that there exist more than one vehicle within each interval of length dy can be ignored and the probability that there exists exactly one vehicle within dy is ρdy. The probability that there exists a vehicle in $[y,y+dy]$ which is also in K_1 is then given by $\rho p_1(y)dy$. Note that the vehicles at x and y are directly connected to each other with probability $g_v^{\mathcal{C}}(\|x-y\|)$. Therefore, the probability that a vehicle at x is directly connected to a vehicle in K_1 and is located in $[y,y+dy]$ is $g_v^{\mathcal{C}}(\|x-y\|)\rho p_1(y)dy$.

Let $h(x,y)$ denotes the probability that the vehicle at x is *not* directly connected to any of the vehicles in K_1 located within $[0,y]$. Because the events that distinct pairs of vehicles are directly connected are independent, the event that the vehicle at x is *not* directly connected to any of the vehicles in K_1 located within $[0,y]$ is independent of the event that the same vehicle is not directly connected to the vehicle in K_1 located within $[y,y+dy]$ (if there is any). We have

$$h(x,y+dy)=h(x,y)(1-g_v^{\mathcal{C}}(\|x-y\|)\rho p_1(y)dy) \tag{9}$$

where the second term on the right hand side of the equation is the complement of the probability that a vehicle at x is directly connected to a vehicle in K_1 and located in $[y,y+dy]$. Eq. (9) leads to

$$dh(x,y)=-h(x,y)g_v^{\mathcal{C}}(\|x-y\|)\rho p_1(y)dy. \tag{10}$$

Therefore the probability that a vehicle at x is not directly connected to any vehicle in K_1 is

$$h(x)=e^{-\int_0^L g_v^{\mathcal{C}}(\|x-y\|)\rho p_1(y)dy}. \tag{11}$$

The result follows immediately. ■

The following two theorems give the access probability p_a and the connectivity probability p_c respectively.

Theorem 1. *Denote by $p_a(x)$ the access probability of a vehicle at x, i.e. the probability that the vehicle at x is connected to either BS1 or BS2 in at most two hops. Then*

$$p_a(x) = 1 - (1 - p_1(x))(1 - p_2(x)) \tag{12}$$

where $p_1(x)$ is given by Eq. (2) *and $p_2(x)$ is given by Eq.* (8).

Proof: The result follows immediately from the observation that the event that a vehicle at x is directly connected to either BS1 or BS2 is independent of the event that the same vehicle is directly connected to at least one vehicle in K_1. ■

Theorem 2 (Approximate result)**.** *Denote by p_c the connectivity probability of $G(L, \rho, \mathcal{C})$, i.e. the probability that all vehicles in the subnetwork $G(L, \rho, \mathcal{C})$ are connected to either BS1 or BS2 in at most two hops. Assume that the event that a vehicle is connected to either BS1 or BS2 in at most two hops is independent of the event that another vehicle is connected to either BS1 or BS2 in at most two hops. Then*

$$p_c = e^{-\int_0^L \rho(1-p_a(x))dx} \tag{13}$$

where $p_a(x)$ is given by Eq. (12).

Proof: Let K_2 be the set of vehicles in $G(L, \rho, \mathcal{C})$ which are connected to either BS1 or BS2 in exactly two hops. Together with the definition of K_1 in Lemma 1, let $\overline{K_1 + K_2} = K \backslash (K_1 + K_2)$ be the set of vehicles in $G(L, \rho, \mathcal{C})$ which are not connected to either BS1 or BS2 in at most two hops. Apply the thinning procedure for K, i.e. consider a realization of K and remove each vehicle located at x independently from this realization, with probability $p_a(x)$. The resulting set of vehicles can be viewed as a realization of $\overline{K_1 + K_2}$ under our assumption that the event that one vehicle is connected to either BS in two hops is independent of the event that another vehicle is connected to either BS in two hops, and the probability that vehicle at x is connected to either BS in two hops is $p_a(x)$. Using the same technique as that used in the proof of Lemma 1, it can be readily shown that $\overline{K_1 + K_2}$ has an inhomogeneous Poisson distribution with density $\rho(1 - p_a(x))$. Then all vehicles $G(L, \rho, \mathcal{C})$ are connected to either BS1 or BS2 in at most two hops if and only if $\mathcal{N}(\overline{K_1 + K_2}) = 0$. The result follows. ■

Note that Theorem 2 only gives an approximate result for the connectivity probability because of the independence assumption. The following lemma proves, in a way, that the event that a vehicle is connected to either BS1 or BS2 in at most two hops is *not* independent of the event that another vehicle is connected to either BS1 or BS2 in at most two hops.

Lemma 3. *Let $h(x) = 1 - p_2(x)$ be the probability that a vehicle at x is not directly connected to any vehicle in K_1; let $h(x_1, x_2)$ be the probability that two vehicles, at x_1 and x_2 respectively, are not directly connected to any vehicle in K_1. Then, $h(x_1, x_2) \geq h(x_1)h(x_2)$.*

Proof: Let $h(x_1, x_2; y)$ denotes the probability that two vehicles, at x_1 and x_2 respectively, are *not* directly connected to any vehicle in K_1 located in $[0, y]$. Using the similar argument in Eq. (9), we have

$$h(x_1, x_2; y + dy) = h(x_1, x_2; y)k(x_1, x_2; y) \tag{14}$$

where $k(x_1, x_2; y) = (1 - g_v^{\mathcal{C}}(\|x_1 - y\|))(1 - g_v^{\mathcal{C}}(\|x_2 - y\|))\rho p_1(y)dy + (1 - \rho p_1(y)dy)$. The first term on the right hand side of $k(x_1, x_2; y)$ is the probability there is a vehicle in K_1 located in $[y, y + dy]$ and both vehicles in x_1 and x_2 are not directly connected to it. The second term on the right hand side of $k(x_1, x_2; y)$ is the probability that there is no vehicle in K_1 located in $[y, y + dy]$. Expanding the right hand side of $k(x_1, x_2; y)$ we have

$$\begin{aligned} k(x_1, x_2; y) = 1 &- g_v^{\mathcal{C}}(\|x_1 - y\|)\rho p_1(y)dy \\ &- g_v^{\mathcal{C}}(\|x_2 - y\|)\rho p_1(y)dy \\ &+ g_v^{\mathcal{C}}(\|x_1 - y\|)g_v^{\mathcal{C}}(\|x_2 - y\|)\rho p_1(y)dy \end{aligned}$$

Using the same approach in Lemma 2 we obtain

$$\begin{aligned} h(x_1, x_2) &= e^{-\int_0^L \left[g_v^{\mathcal{C}}(\|x_1 - y\|) + g_v^{\mathcal{C}}(\|x_2 - y\|)\right]\rho p_1(y)dy} \\ &\quad \times e^{\int_0^L g_v^{\mathcal{C}}(\|x_1 - y\|)g_v^{\mathcal{C}}(\|x_2 - y\|)\rho p_1(y)dy} \\ &\geq e^{-\int_0^L \left[g_v^{\mathcal{C}}(\|x_1 - y\|) + g_v^{\mathcal{C}}(\|x_2 - y\|)\right]\rho p_1(y)dy} \\ &= h(x_1)h(x_2) \qquad \text{(from Eq. (11))} \end{aligned}$$

■

Before obtaining the exact result of the connectivity probability, we introduce some properties in the following lemma.

Lemma 4. *Let $p_c(\mathbf{y})$ be the connectivity probability of $G(L, \rho, \mathcal{C})$ conditioned on that the number of vehicles directly connected to either BS is n and they are located at $\mathbf{y} = \{y_1, y_2, \cdots, y_n : 0 \leq y_i \leq L, 1 \leq i \leq n\}$; let $p_Y(\mathbf{y})$ be the probability density function (pdf) of $\mathbf{y}$ conditioned on that there are n vehicles directly connected to either BS. The following properties hold.*

$$\textit{(i)} \quad p_Y(\mathbf{y}) = \prod_{i=1}^{n} \frac{p_1(y_i)}{\int_0^L p_1(x)dx} \tag{15}$$

$$\textit{(ii)} \quad p_c(\mathbf{y}) = e^{-\int_0^L \rho(1-p_1(x)) \prod_{i=1}^{n}(1 - g_v^{\mathcal{C}}(\|x - y_i\|))dx} \tag{16}$$

Proof: For $n = 1$, $p_Y(y_1) = \frac{p_1(y_1)}{\int_0^L p_1(x)dx}$ is the probability that a vehicle in K_1 is located at y_1. Since $p_1(y_i)$ and $p_1(y_j)$ are mutually independent for $i \neq j$, the result follows for Eq. (15).

For Eq. (16), note that a vehicle at x is not connected to any BSs in at most two hops if it is not directly connected to any BSs (the probability is $1 - p_1(x)$) *and* it is not directly connected to vehicles which are located at $\mathbf{y}$ given that these vehicles are in K_1 (the probability is $1 - g_v^{\mathcal{C}}(\|x - y_i\|)$ for $1 \leq i \leq n$). That is, vehicle at x cannot access any BS in at most two hops with probability

$$(1 - p_1(x)) \prod_{i=1}^{n} (1 - g_v^{\mathcal{C}}(\|x - y_i\|)). \tag{17}$$

Eq. (17) is valid when $x \notin \mathbf{y}$. When $x = y_j$ for arbitrary j, we assume that $g_v^{\mathcal{C}}(0) = 1$. This implies that $p_a(x|\mathbf{y}) = 1$. So Eq. (17) is still valid when $x \in \mathbf{y}$. Applying the thinning

procedure and the technique used in Lemma 1, we have the number of vehicles which are neither directly connected to any BSs nor directly connected to any of the vehicles at $\mathbf{y}$ is an inhomogeneous Poisson random variable with density $\rho(1-p_1(x))\prod_{i=1}^{n}(1-g_v^{\mathcal{C}}(\|x-y_i\|))$. The result follows immediately. ∎

Theorem 3 (Exact result)**.** *Denote by p_c the connectivity probability of $G(L,\rho,\mathcal{C})$, i.e. the probability that all vehicle in the subnetwork $G(L,\rho,\mathcal{C})$ is connected to either BS1 or BS2 in at most two hops. Then*

$$p_c=\sum_{n=0}^{\infty}\Pr(\mathcal{N}(K_1)=n)\left[\int_{[0,L]^n}p_c(\mathbf{y})p_Y(\mathbf{y})d\mathbf{y}\right] \quad (18)$$

where $p_c(\mathbf{y})$ and $p_Y(\mathbf{y})$ are given by Lemma 4; $\Pr(\mathcal{N}(K_1)=n)$ *is given by Lemma 1. When $n=0$, we declare*

$$\int_{[0,L]^n}p_c(\mathbf{y})p_Y(\mathbf{y})d\mathbf{y}\Big|_{n=0}=p_c(\mathbf{y})p_Y(\mathbf{y})\Big|_{n=0}$$
$$=e^{-\int_0^L\rho(1-p_1(x))dx}.$$

Proof: Eq. (18) directly follows from the law of total probability, so the details are omitted here. ∎

Eq. (18) gives an exact formula for the connectivity probability which does not rely on the assumption that the event that a vehicle is connected to either BS in two hops and the event that another vehicle is connected to either BS in two hops are independent. However Eq. (18) is much more complicated than the approximate result in Eq. (13). In many situations, Eq. (13) provides a reasonably accurate result for the connectivity probability. Therefore we include both results in this paper.

V. Performance Evaluation under Specific Channel Models

Based on the analysis in Section IV, we further derive and compare in this section the access probability and connectivity probability performance under two specific channel models, i.e. unit disk model and log-normal shadowing model.

A. Unit Disk Model

In the unit disk model $\mathcal{U}$, assume that two vehicles are directly connected if and only if their Euclidean distance is less than or equal to r; assume that a vehicle and a BS are directly connected if and only if their Euclidean distance is not more than R. In other words,

$$g_v^{\mathcal{U}}(x)=\begin{cases}1 & \text{if } x\le r\\ 0 & \text{otherwise,}\end{cases}\qquad g_b^{\mathcal{U}}(x)=\begin{cases}1 & \text{if } x\le R\\ 0 & \text{otherwise.}\end{cases}$$

where r and R are predetermined values, commonly known as the transmission ranges. Typically we have $R>r$. Applying the above equations into Eq. (2), (8) and (12) we obtain the access probability under the unit disk model $\mathcal{U}$:

(I) For $0<L\le 2R$,
we have $p_1(x)=1$ implies that $p_a(x)=1$ for $x\in[0,L]$. Hence,

$$p_a=1.$$

(II) For $2R<L\le 2R+r$,
we have $p_1(x)$ is 0 when $x\in(R,L-R)$, and 1 otherwise. When $x\in(R,L-R)$, Eq. (8) becomes

$$p_2(x)=1-e^{-\int_0^R g_v^{\mathcal{C}}(\|x-y\|)\rho dy-\int_{L-R}^{L}g_v^{\mathcal{C}}(\|x-y\|)\rho dy}$$
$$=1-e^{-\int_{x-r}^{R}\rho dy-\int_{L-R}^{x+r}\rho dy}$$
$$=1-e^{-\rho(2R+2r-L)}.$$

So substitute $p_2(x)$ into Eq. (12),

$$p_a=\frac{2R}{L}+\frac{L-2R}{L}(1-e^{-\rho(2R+2r-L)})$$
$$=1-\frac{L-2R}{L}e^{-\rho(2R+2r-L)}.$$

(III) For $2R+r<L\le 2R+2r$,
we have for $x\in(R,L-R)$, Eq. (8) becomes

$$p_2(x)=1-e^{-\int_0^R g_v^{\mathcal{C}}(\|x-y\|)\rho dy-\int_{L-R}^{L}g_v^{\mathcal{C}}(\|x-y\|)\rho dy}$$
$$=1-e^{-\int_{x-r}^{R}\rho dy-\int_{L-R}^{x+r}\rho dy}.$$

So substitute $p_2(x)$ into Eq. (12),

$$p_a=\frac{2R}{L}+\frac{1}{L}\int_R^{L-R-r}(1-e^{-\rho(R+r-x)})dx$$
$$+\frac{1}{L}\int_{L-R-r}^{R+r}(1-e^{-\rho(2R+2r-L)})dx$$
$$+\frac{1}{L}\int_{R+r}^{L-R}(1-e^{-\rho(R+r+x-L)})dx$$
$$=1+\frac{2}{\rho L}(e^{-\rho r}-e^{-\rho(2R+2r-L)})$$
$$-\frac{2R+2r-L}{L}e^{-\rho(2R+2r-L)}.$$

(IV) For $L>2R+2r$,
From Eq. (8)

$$p_2(x)=\begin{cases}1-e^{-\int_{x-r}^{R}\rho dy}=1-e^{-\rho(R+r-x)}\\ \qquad\text{when } x\in(R,R+r],\\ 1-e^{-\int_{L-R}^{x+r}\rho dy}=1-e^{-\rho(R+r+x-L)}\\ \qquad\text{when } x\in[L-R-r,L-R),\\ 0\qquad\text{when } x\in(R+r,L-R-r).\end{cases}$$

So substituting $p_2(x)$ into Eq. (12),

$$p_a=\frac{2R}{L}+\frac{1}{L}\int_R^{R+r}(1-e^{-\rho(R+r-x)})dx$$
$$+\frac{1}{L}\int_{L-R-r}^{L-R}(1-e^{-\rho(R+r+x-L)})dx$$
$$=\frac{2R+2r}{L}+\frac{2(e^{-\rho r}-1)}{\rho L}.$$

To derive the equations for the connectivity probability (exact result), we first look at Lemma 4. For the unit disk model, $p_1(x)$ is 1 when $x\in[0,R]\cup[L-R,L]$ and zero otherwise. Hence, Eq. (15) becomes

$$p_Y(\mathbf{y})=\frac{1}{(\min(2R,L))^n} \quad (19)$$

when $y_i \in [0,R] \cup [L-R,L], \forall y_i$, and zero otherwise. Eq. (16) becomes

$$p_c(\mathbf{y}) = e^{-\int_R^{L-R} \rho \prod_{i=1}^{n} (1-g_v^{\mathcal{C}}(\|x-y_i\|))dx}. \quad (20)$$

Note that $\prod_{i=1}^{n}(1-g_v^{\mathcal{C}}(\|x-y_i\|))$ is 1 when $\|x-y_i\| > r$ for all y_i. For $L \le 2R$, we can easily obtain p_c from Eq. (18) by substituting Eq. (19) and (20) into it (will be shown later). To obtain the result for $L > 2R$, the following transformation will simplify the arithmetic work.

Let S_a (and S_b) be the set of vehicles in $[0,R]$ (and $[L-R,L]$) which, by definition, are also in K_1. Let $\mathcal{N}(S_a)$ (and $\mathcal{N}(S_b)$) be the number of vehicles in S_a (and S_b). Note that $S_a \cup S_b = K_1$ and $\mathcal{N}(S_a)+\mathcal{N}(S_b) = \mathcal{N}(K_1)$. Let y_a (and y_b) be the location of the vehicle in S_a (and S_b) which is furthest from BS1 (and BS2). That is,

$$y_a = \begin{cases} 0 & \text{if } S_a = \emptyset \\ \max\{y : y \in S_a\} & \text{otherwise,} \end{cases} \quad (21)$$

$$y_b = \begin{cases} L & \text{if } S_b = \emptyset \\ \min\{y : y \in S_b\} & \text{otherwise.} \end{cases} \quad (22)$$

Therefore, the cumulative probability function of y_a is

$$\begin{aligned} \Pr(y_a \le y_{max}) &= \Pr(y_i \le y_{max}, \forall y_i \in S_a) \\ &= (\frac{y_{max}}{R})^{n_a} \qquad \text{for } n_a = \mathcal{N}(S_a) \ge 1. \end{aligned}$$

With Eq. (21) defines $y_a = 0$ when $\mathcal{N}(S_a) = 0$, we have the pdf of y_a as

$$f_a(y_a; n_a) = \begin{cases} \frac{n_a}{R}(\frac{y_a}{R})^{n_a-1} & \text{if } n_a \ge 1 \\ \delta(y_a) & \text{if } n_a = 0. \end{cases}$$

Similarly we have the pdf of y_b as

$$f_b(y_b; n_b) = \begin{cases} \frac{n_b}{R}(\frac{L-y_b}{R})^{n_b-1} & \text{if } n_b \ge 1 \\ \delta(L-y_b) & \text{if } n_b = 0. \end{cases}$$

With y_a and y_b, we can rewrite Eq. (20) into

$$p_c(y_a, y_b) = e^{-\int_{\max\{R, y_a+r\}}^{\min\{L-R, y_b-r\}} \rho dx}, \quad (23)$$

and Eq. (18) can be transformed into

$$\begin{aligned} p_c = \sum_{n_a=0}^{\infty} \sum_{n_b=0}^{\infty} &\Pr(\mathcal{N}(S_a) = n_a) \Pr(\mathcal{N}(S_b) = n_b) \\ &\left[\int_0^R \int_{L-R}^{L} p_c(y_a, y_b) f_a(y_a; n_a) f_b(y_b; n_b) dy_b dy_a \right] \end{aligned} \quad (24)$$

for $L > 2R$. Eq. (24) can be further simplified under different cases. For $n_a > 0$ and $n_b > 0$, Eq. (24) becomes

$$\begin{aligned} & p_c^{(n_a>0, n_b>0)} \\ &= \int_0^R \int_{L-R}^{L} p_c(y_a, y_b) \left[\sum_{n_a=1}^{\infty} \sum_{n_b=1}^{\infty} \Pr(\mathcal{N}(S_a) = n_a) \right. \\ &\quad \left. \Pr(\mathcal{N}(S_b) = n_b) f_a(y_a; n_a) f_b(y_b; n_b) \right] dy_b dy_a \\ &= \int_0^R \int_{L-R}^{L} p_c(y_a, y_b) \left[\sum_{n_a=1}^{\infty} \sum_{n_b=1}^{\infty} \frac{(\rho R)^{n_a}}{n_a!} e^{-\rho R} \right. \\ &\quad \left. \frac{(\rho R)^{n_b}}{n_b!} e^{-\rho R} \frac{n_a}{R} (\frac{y_a}{R})^{n_a-1} \frac{n_b}{R} (\frac{L-y_b}{R})^{n_b-1} \right] dy_b dy_a \\ &= \int_0^R \int_{L-R}^{L} p_c(y_a, y_b) \rho^2 e^{-2\rho R} \\ &\quad \left[\sum_{n_a=1}^{\infty} \frac{(\rho y_a)^{n_a-1}}{(n_a-1)!} \right] \left[\sum_{n_b=1}^{\infty} \frac{(\rho(L-y_b))^{n_b-1}}{(n_b-1)!} \right] dy_b dy_a \\ &= \int_0^R \int_{L-R}^{L} p_c(y_a, y_b) \rho^2 e^{-2\rho R} e^{\rho y_a} e^{\rho(L-y_b)} dy_b dy_a. \end{aligned} \quad (25)$$

For $n_a = 0$ and $n_b > 0$, Eq. (24) becomes

$$\begin{aligned} & p_c^{(n_a=0, n_b>0)} \\ &= \sum_{n_b=1}^{\infty} e^{-\rho R} \frac{(\rho R)^{n_b}}{n_b!} e^{-\rho R} \left[\int_{L-R}^{L} p_c(0, y_b) \frac{n_b}{R} (\frac{L-y_b}{R})^{n_b-1} dy_b \right] \\ &= \int_{L-R}^{L} p_c(0, y_b) \rho e^{-2\rho R} \sum_{n_b=1}^{\infty} \frac{(\rho(L-y_b))^{n_b-1}}{(n_b-1)!} dy_b \\ &= \int_{L-R}^{L} p_c(0, y_b) \rho e^{-2\rho R} e^{\rho(L-y_b)} dy_b. \end{aligned} \quad (26)$$

With similar steps (omit here) we can obtain for $n_a > 0$ and $n_b = 0$, Eq. (24) becomes

$$p_c^{(n_a>0, n_b=0)} = \int_0^R p_c(y_a, L) \rho e^{-2\rho R} e^{\rho y_a} dy_a. \quad (27)$$

Note that it can be shown that Eq. (27) equals to (26) by letting $y_b = L - y_a$, then

$$\begin{aligned} p_c^{(n_a>0, n_b=0)} &= -\int_{L}^{L-R} p_c(L-y_b, L) \rho e^{-2\rho R} e^{\rho(L-y_b)} dy_b \\ &= \int_{L-R}^{L} p_c(0, y_b) \rho e^{-2\rho R} e^{\rho(L-y_b)} dy_b \end{aligned}$$

where $p_c(L-y_b, L) = p_c(0, y_b)$. Finally for $n_a = 0$ and $n_b = 0$,

$$\begin{aligned} p_c^{(n_a=0, n_b=0)} &= e^{-\rho R} e^{-\rho R} p_c(0, L) = e^{-2\rho R} e^{-\int_R^{L-R} \rho dx} \\ &= e^{-2\rho R} e^{-\rho(L-2R)} = e^{-\rho L}. \end{aligned} \quad (28)$$

Using Eq. (25), (26), (27) and (28), we can obtain the connectivity probability as follows. Due to the lengthy (but straightforward) steps involved to derive the results, we omit the intermediate steps and only include the results of Eq. (25) and (26) for readers' convenience.

(I) For $0 < L \le 2R$,
$p_Y(\mathbf{y}) = \frac{1}{L^n}$ from Eq. (15) and $p_c(\mathbf{y}) = 1$ from Eq. (16)

implies that

$$p_c = \sum_{n=0}^{\infty} \Pr(\mathcal{N}(K_1) = n) = 1.$$

(II) For $2R < L \leq 2R + r$,

$$\begin{aligned} p_c^{(n_a>0,n_b>0)} &= 1 + e^{-\rho L} - 2e^{-\rho R} + e^{-\rho(3R+r-L)} \\ &\quad + (-\frac{1}{4} - \frac{1}{2}\rho(L-2R))e^{-\rho(2R+2r-L)} \\ &\quad - e^{-\rho(L+r-R)} + \frac{1}{4}e^{-\rho(L+2r-2R)}, \\ p_c^{(n_a=0,n_b>0)} &= -e^{-\rho L} + e^{-\rho R} - \frac{1}{2}e^{-\rho(3R+r-L)} \\ &\quad + \frac{1}{2}e^{\rho(L+r-R)}, \\ p_c &= 1 + \frac{1}{4}e^{-\rho(L+2r-2R)} \\ &\quad + (-\frac{1}{4} - \frac{1}{2}\rho(L-2R))e^{-\rho(2R+2r-L)}. \end{aligned}$$

(III) For $2R + r < L \leq 2R + 2r$,

$$\begin{aligned} & p_c^{(n_a>0,n_b>0)} \\ &= 1 + e^{-\rho L} + \frac{1}{2}e^{-\rho(L-2R)} - e^{-\rho(L+r-R)} \\ &\quad + (-\frac{3}{4} - \frac{1}{2}\rho(2R+2r-L))e^{-\rho(2R+2r-L)} \\ &\quad + \frac{1}{4}e^{-\rho(L+2r-2R)} - e^{-\rho(L-R-r)}, \\ & p_c^{(n_a=0,n_b>0)} \\ &= -e^{-\rho L} + \frac{1}{2}e^{-\rho(L+r-R)} + \frac{1}{2}e^{-\rho(L-R-r)}, \\ p_c &= 1 + \frac{1}{2}e^{-\rho(L-2R)} + \frac{1}{4}e^{-\rho(L+2r-2R)} \\ &\quad + (-\frac{3}{4} - \frac{1}{2}\rho(2R+2r-L))e^{-\rho(2R+2r-L)}. \end{aligned}$$

(IV) For $L > 2R + 2r$,

$$\begin{aligned} & p_c^{(n_a>0,n_b>0)} \\ &= e^{-\rho L} + \frac{1}{4}e^{-\rho(L+2r-2R)} - e^{-\rho(L+r-R)} \\ &\quad + \frac{1}{4}e^{-\rho(L-2R-2r)} - e^{-\rho(L-R-r)} + \frac{1}{2}e^{-\rho(L-2R)}, \\ & p_c^{(n_a=0,n_b>0)} \\ &= -e^{-\rho L} + \frac{1}{2}e^{-\rho(L+r-R)} + \frac{1}{2}e^{-\rho(L-R-r)}, \\ p_c &= \frac{1}{4}e^{-\rho(L+2r-2R)} + \frac{1}{4}e^{-\rho(L-2R-2r)} + \frac{1}{2}e^{-\rho(L-2R)}. \end{aligned}$$

B. Log-normal Shadowing Model

The log-normal model $\mathcal{L}$ is commonly used to model the real world signal propagation where the transmit power loss increases logarithmically with the Euclidean distance between two wireless nodes and varies log-normally due to the shadowing effect caused by surrounding environment. In the log-normal model, we formulate the received power (in dB) at a destination vehicle as

$$p_{rx} = p_0 - 10\alpha \log_{10} \frac{l}{d_0} + N_\sigma \tag{29}$$

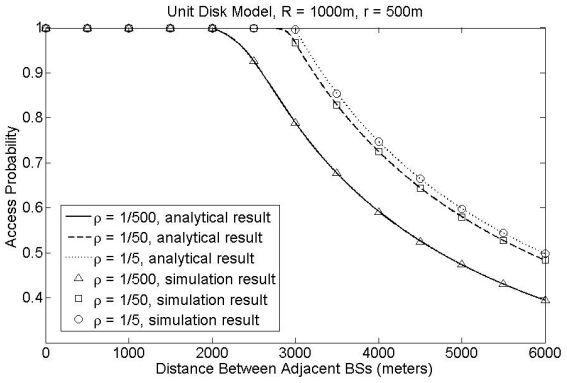


Fig. 3. Access probability with L changing under the unit disk model, $R = 1000$m, $r = 500$m, $\rho = 1/5, 1/50, 1/500$ vehicles/m respectively.

where p_{rx} is the received power (in dBmW) at the destination vehicle; p_0 is the power (in dBmW) at a reference distance d_0; α is the path loss exponent; N_σ is a Gaussian random variable with zero mean and variance σ^2; l is the Euclidean distance between the two vehicles (or a vehicle and a BS depending on the context). A source vehicle can establish a direct connection to a destination vehicle if the received power at the destination vehicle p_{rx} is greater than or equal to a certain threshold power p_{th}^v. Similarly, a source vehicle can establish a two-way direct connection to a destination BS if the received power at the destination BS p_{rx} is greater than or equal to a certain threshold power p_{th}^b. In this paper, we assume that wireless connections between vehicles, and between vehicles and BSs, are symmetric. Note that when $\sigma = 0$, the log-normal model reduces to the unit disk model. Due to this fact, we assign $p_{th}^v = p_0 - 10\alpha \log_{10} \frac{r}{d_0}$, $p_{th}^b = p_0 - 10\alpha \log_{10} \frac{R}{d_0}$ so that the results under log-normal model can be compared with the results under the unit disk model later. It can be shown that under the log-normal model

$$g_v^{\mathcal{L}}(x) = \Pr(p_{rx} \geq p_{th}^v) = Q(\frac{10\alpha}{\sigma} \log_{10} \frac{x}{r}).$$

where function $Q(y) = \frac{1}{\sqrt{2\pi}} \int_y^\infty e^{-\frac{x^2}{2}} dx$ is the tail probability of the standard normal distribution. Similarly, $g_b^{\mathcal{L}}(x) = Q(\frac{10\alpha}{\sigma} \log_{10} \frac{x}{R})$. When $\sigma = 0$, $g_v^{\mathcal{L}}(x) = \Pr(x \leq r)$, $g_b^{\mathcal{L}}(x) = \Pr(x \leq R)$ and the log-normal model becomes the unit disk model as expected.

The access probability can then be obtained for different values of α and σ by computing Eq. (12) using any numerical integration technique. The approximate and exact results for the connectivity probability can be obtained by computing Eq. (13) and (18) using any numerical integration technique.

VI. Analytical and Simulation Results

A. Unit Disk Model

Fig. 3 shows the access probability given different values of L and ρ. The analytical results are verified by the simulation results which are obtained from 40000 randomly generated network topologies. As the number of instances of random networks used in the simulation is very large, the confidence interval is too small to be distinguishable and hence ignored in this plot as well as other plots. As shown in the figure, the access probability decreases with L when L exceeds some

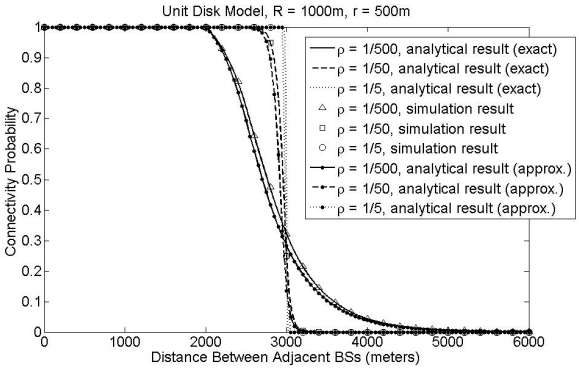


Fig. 4. Connectivity probability with L changing under the unit disk model, $R = 1000$m, $r = 500$m, $\rho = 1/5, 1/50, 1/500$ vehicles/m respectively.

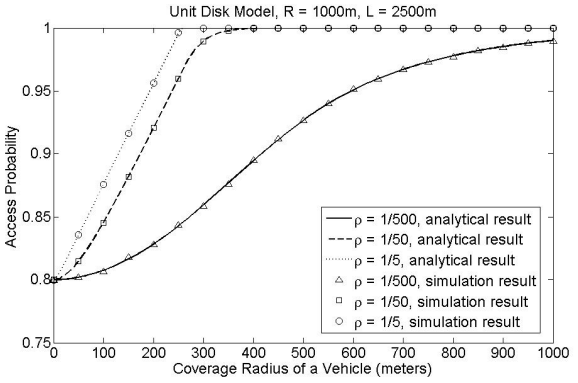


Fig. 5. Access probability with r changing under the unit disk model, $R = 1000$m, $L = 2500$m, $\rho = 1/5, 1/50, 1/500$ vehicles/m respectively.

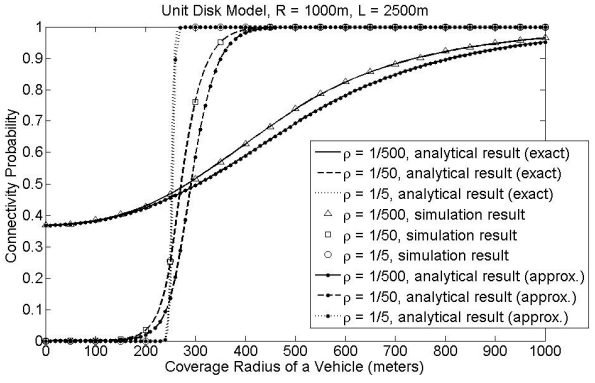


Fig. 6. Connectivity probability with r changing under the unit disk model, $R = 1000$m, $L = 2500$m, $\rho = 1/5, 1/50, 1/500$ vehicles/m respectively.

limits. For small ρ, the access probability decreases as soon as $L > 2R$. That is because when the vehicle density ρ (number of vehicles per meter) is low, a vehicle is either directly connected to a BS or disconnected, i.e cannot reach any BS in at most two hops. It is hard for the vehicle to find a one-hop relay in its range via which it can access a BS if it is not within the transmission range of any BS. However for large ρ, it is easier for the vehicle, which is not within the transmission range of any BS, to find a one-hop relay to access the BS. In general the access probability increases with an increase in ρ, and the reason is that when the vehicle density increases, the probability increases for vehicles in the gap of the transmission ranges of BSs to find a neighbor within the transmission range of a BS to act as a relay.

Similarly, Fig. 4 shows the connectivity probability for different values of L and ρ. The exact analytical results are verified by the simulation results. The approximate analytical result is shown to be reasonably close to the exact analytical result. The figure shows that when $L \leq 2R+r = 2500$ meters, it is easy for all vehicles to be connected to either BS in at most two hops, hence the connectivity probability is high. As L gets larger, it is harder for all vehicles to be connected to the BSs due to the larger possible distances between the vehicles and the BSs. This causes a drop in the connectivity probability, and the connectivity probability tends to zero as L goes to infinity. The transition of the connectivity probability from 1 to 0 gets sharper as the vehicle density increases. As ρ goes to infinity, the transition happens at the critical distance $L = 2R + 2r = 3000$ meters, below which the network is disconnected with a high probability and above which the network is connected with a high probability. Furthermore, the networks with a larger ρ have a higher connectivity probability than the networks with a smaller ρ when L is small. This is because when the vehicle density is large, it is easier for vehicles not directly connected to a BS to find a vehicle within its communication range and is directly connected to a BS to act as a relay. When L is large, the networks with a larger ρ have a lower connectivity probability than the networks with a smaller ρ. This is because at large values of L when the vehicle density is large it is easier to have at least one vehicle which is located too far from the BSs to be connected to a BS in at most two hops.

Fig. 5 shows how the transmission range of the vehicles r affect the access probability. It shows that the access probability increases with r, and when ρ is large enough, the access probability could be quite close to 1. And it shows again that the access probability increases with an increase in ρ.

With a similar setup, Fig. 6 shows the sensitivity of the connectivity probability to r. For a large ρ, around a certain value of r a small increase in r will incur a dramatic increase in the connectivity probability from near 0 to near 1, i.e. the well-known phase transition phenomenon. From the figure it shows that such phenomenon does not exist for small ρ. Fig. 6 also shows a scenario where there may be a significant gap between the approximate and exact results for connectivity probability.

Fig. 7 supported our conclusion that an increase in ρ will improve the access probability as it shows that the access probability monotonically increases with ρ. While ρ is relatively small, and the width of the gap region not directly covered by any of the BSs is relatively large, the access probability will be low, and thus, in this circumstance, network operator should consider to deploy more BSs along the highway for better connectivity and greater access probability.

B. Log-normal Shadowing Model

Fig. 8 shows the access probability under the log-normal shadowing model. In general, it is easier for the vehicles in the subnetwork to get access to any BS under the log-normal model. As σ increases, the access probability improves. The improvement in access probability is more significant for high vehicular density.

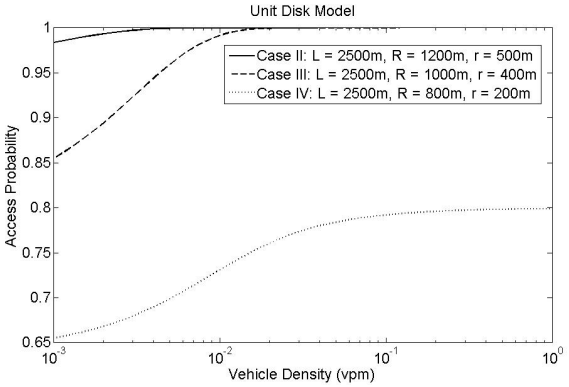


Fig. 7. Access probability with ρ changing under the unit disk model, L, R, r are constants.

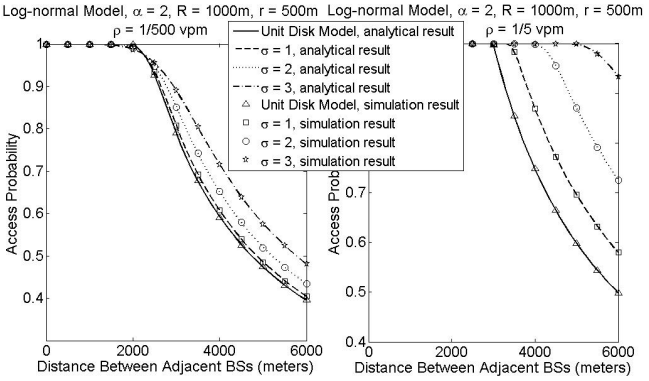


Fig. 8. Access probability with L changing under the log-normal model, $R = 1000$m, $r = 500$m, $\rho = 1/5, 1/500$ vehicles/m respectively under different values of σ. R (r) is the transmission range of a BS (vehicle) ignoring shadowing effect, i.e. $\sigma = 0$.

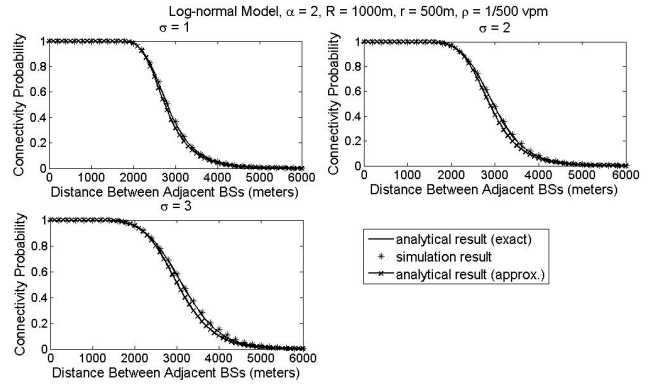


Fig. 9. Connectivity probability with L changing under the log-normal model, $R = 1000$m, $r = 500$m, $\rho = 1/500$ vehicles/m under different values of σ. R (r) is the transmission range of a BS (vehicle) ignoring shadowing effect, i.e. $\sigma = 0$.

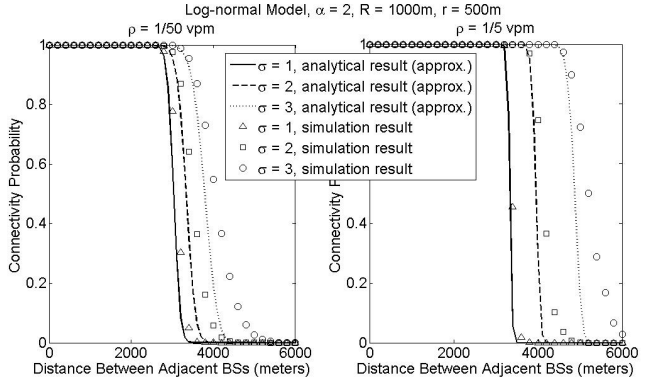


Fig. 10. Connectivity probability with L changing under the log-normal model, $R = 1000$m, $r = 500$m, $\rho = 1/5, 1/50$ vehicles/m respectively under different values of σ. R (r) is the transmission range of a BS (vehicle) ignoring shadowing effect, i.e. $\sigma = 0$.

Fig. 9 shows the connectivity probability under the log-normal model when the vehicle density is low ($\rho = \frac{1}{500}$ vpm). The exact analytical results are verified by the simulation results. As the vehicle density increases, the computational complexity involved in numerically computing the exact result increases very quickly. As such, we only provide the exact analytical results for low vehicle density. Furthermore, Fig. 9 shows that the approximate analytical results are reasonably close to the true values when the vehicle density is low. However, as shown in Fig. 10, the discrepancy between the approximate results and the true values can be significant when the vehicle density is high ($\rho = \frac{1}{50}, \frac{1}{5}$ vpm). In general, the approximate analytical result always under-estimate the simulation result. Same situation can be observed for the result under the unit disk model. This can be explained by Lemma 3 that a vehicle is more likely to be able to access to any BS where there is another vehicle nearby that can access to the BSs. Because of the independence assumption used in obtaining the approximate analytical result, the approximate result will under-estimate the true value.

VII. Conclusions

In this paper, we analyzed the connectivity probability and the access probability for a given network bounded by two adjacent base stations, and vehicles in the network are Poissonly distributed with known density and each vehicle can communicate with a base station in at most two hops. Under a general connection model, and later on taking the unit disk communication model and the log-normal shadowing model as the specific examples, we derived closed-form formulas for the access probability and connectivity probability considering that the base stations and the vehicles have different transmission capabilities. These formulas characterize the relation between these key parameters, i.e. the transmission ranges of the base stations and the vehicles, the distance between adjacent base stations, the vehicle density and their impact on the access and connectivity probabilities. These results can be useful for a network operator to design a network with a given level of access guarantee. In future, we plan to extend the current work on 1-D networks to 2-D networks.

References

[1] S. Sai, E. Niwa, K. Mase, M. Nishibori, J. Inoue, M. Obuchi, T. Harada, H. Ito, K. Mizutani, and M. Kizu, "Field evaluation of UHF radio propagation for an ITS safety system in an urban environment," *IEEE Commun. Mag.*, vol. 47, no. 11, pp. 120–127, 2009.

[2] C.-X. Wang, X. Hong, X. Ge, X. Cheng, G. Zhang, and J. S. Thompson, "Cooperative mimo channel models: A survey," *IEEE Commun. Mag.*, vol. 48, no. 2, pp. 80–87, 2010.

[3] C.-X. Wang, X. Cheng, and D. I. Laurenson, "Vehicle-to-vehicle channel modeling and measurements: recent advances and future challenges," *IEEE Commun. Mag.*, vol. 47, no. 11, pp. 96–103, 2009.

[4] "IEEE P802.11 - task group p," http://grouper.ieee.org/groups/802/11/Reports/tgp_update.htm.

[5] "IEEE 1609 - family of standards for wireless access in vehicular environments (WAVE)," http://www.standards.its.dot.gov/fact_sheet.asp?f=80, 2006.
[6] G. Acosta-Marum and M. A. Ingram, "Six time- and frequency- selective empirical channel models for vehicular wireless LANs," *IEEE Veh. Technol. Mag.*, vol. 2, no. 4, pp. 4 – 11, 2007.
[7] Y. Chang, M. Lee, and J. A. Copeland, "An adaptive on-demand channel estimation for vehicular ad hoc networks," in *Consumer Communications and Networking Conference (CCNC)*, 2009, pp. 1–5.
[8] L. Cheng, B. E. Henty, R. Cooper, D. D. Stancil, and D. F. Bai, "A measurement study of time-scaled 802.11a waveforms over the mobile-to-mobile vehicular channel at 5.9 GHz," *IEEE Commun. Mag.*, vol. 46, no. 5, pp. 84–91, 2008.
[9] L. Stibor, Y. Zang, and H.-J. Reumerman, "Evaluation of communication distance of broadcast messages in a vehicular ad-hoc network using IEEE 802.11p," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2007, pp. 254 – 257.
[10] M. Wellens, B. Westphal, and P. Mahonen, "Performance evaluation of IEEE 802.11-based WLANs in vehicular scenarios," in *IEEE Vehicular Technology Conference (VTC)*, 2007, pp. 1167–1171.
[11] S. Eichler, "Performance evaluation of the IEEE 802.11p WAVE communication standard," in *IEEE Vehicular Technology Conference (VTC)*, 2007, pp. 2199–2203.
[12] J. C. Burguillo-Rial, E. Costa-Montenegro, F. Gil-Castineira, and P. Rodriguez-Hernandez, "Performance analysis of IEEE 802.11p in urban environments using a multi-agent model," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2008, pp. 1–6.
[13] A. Vinel, V. Vishnevsky, and Y. Koucheryavy, "A simple analytical model for the periodic broadcasting in vehicular ad-hoc networks," in *IEEE Global Communications Conference (GLOBECOM) Workshops*, 2008, pp. 1–5.
[14] A. Sebastian, M. Tang, Y. Feng, and M. Looi, "Multi-vehicles interaction graph model for cooperative collision warning system," in *IEEE Intelligent Vehicles Symposium*, 2009, pp. 929–934.
[15] Y. Wang, A. Ahmed, B. Krishnamachari, and K. Psounis, "IEEE 802.11p performance evaluation and protocol enhancement," in *IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, 2008, pp. 317–322.
[16] K. Bilstrup, E. Uhlemann, E. G. Strom, and U. Bilstrup, "Evaluation of the IEEE 802.11p MAC method for vehicle-to-vehicle communication," in *IEEE Vehicular Technology Conference (VTC)*, 2008, pp. 1–5.
[17] J. R. Gallardo, D. Makrakis, and H. T. Mouftah, "Performance analysis of the EDCA medium access mechanism over the control channel of an IEEE 802.11p WAVE vehicular network," in *IEEE International Conference on Communications (ICC)*, 2009, pp. 1–6.
[18] N. Ferreira, J. A. Fonseca, and J. S. Gomes, "On the adequacy of 802.11p MAC protocols to support safety services in ITS," in *IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2008, pp. 1189–1192.
[19] M. Amadeo, C. Campolo, A. Molinaro, and G. Ruggeri, "A WAVE-compliant MAC protocol to support vehicle-to-infrastructure non-safety applications," in *IEEE International Conference on Communications (ICC) Workshops*, 2009, pp. 1–6.
[20] S. Y. Wang and C. L. Chou, "NCTUns 5.0: A network simulator for IEEE 802.11(p) and 1609 wireless vehicular network researches," in *IEEE Vehicular Technology Conference (VTC)*, 2008, pp. 1–2.
[21] T. M. Fernandez-Carames, J. A. Garcia-Naya, M. Gonzalez-Lopez, and L. Castedo, "FlexVehd: A flexible testbed for vehicular radio interfaces," in *International Conference on ITS Telecommunications (ITST)*, 2008, pp. 283–287.
[22] J. Peng and L. Cheng, "A distributed MAC scheme for emergency message dissemination in vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 6, pp. 3300 – 3308, 2007.
[23] I. Salhi, M. O. Cherif, and S. M. Senouci, "A new architecture for data collection in vehicular networks," in *IEEE International Conference on Communications (ICC)*, 2009, pp. 1–6.
[24] M. Jonsson and A. Bohm, "Position-based data traffic prioritization in safety-critical, real-time vehicle-to-infrastructure communication," in *IEEE International Conference on Communications (ICC) Workshops*, 2009, pp. 1–6.
[25] M. Torrent-Moreno, J. Mittag, P. Santi, and H. Hartenstein, "Vehicle-to-vehicle communication: Fair transmit power control for safety-critical information," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3684 – 3703, 2009.
[26] B. Shrestha, D. Niyato, H. Zhu, and E. Hossain, "Wireless access in vehicular environments using BitTorrent and bargaining," in *IEEE Global Communications Conference (GLOBECOM)*, 2008, pp. 1–5.
[27] J. Wu, "Connectivity analysis of a mobile vehicular ad hoc network with dynamic node population," in *IEEE Global Communications Conference (GLOBECOM)*, 2008, pp. 1–8.
[28] O. Dousse, P. Thiran, and M. Hasler, "Connectivity in ad-hoc and hybrid networks," in *IEEE Conference on Computer Communications (INFOCOM)*, vol. 2, 2002, pp. 1079–1088.
[29] R. S. Ojha, G. Kannan, S. N. Merchant, and U. B. Desai, "On optimal transmission range for multihop cellular networks," in *IEEE Global Communications Conference (GLOBECOM)*, 2008, pp. 1–5.
[30] S. Mukherjee, D. Avidor, and K. Hartman, "Connectivity, power, and energy in a multihop cellular-packet system," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 818–836, 2007.
[31] X. Ta, G. Mao, and B. D. O. Anderson, "On the connectivity of wireless multi-hop networks with arbitrary wireless channel models," *IEEE Commun. Lett.*, vol. 13, no. 3, pp. 181 – 183, 2009.
[32] P. Agrawal and N. Patwari, "Correlated link shadow fading in multi-hop wireless networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4024–4036, 2009.
[33] C. X. Wang, M. Patzold, and Q. Yao, "Stochastic modeling and simulation of frequency-correlated wideband fading channels," *IEEE Trans. Veh. Technol.*, vol. 56, no. 3, pp. 1050–1063, 2007.
[34] X. Cheng, C.-X. Wang, D. I. Laurenson, S. Salous, and A. V. Vasilakos, "An adaptive geometry-based stochastic model for non-isotropic mimo mobile-to-mobile channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4824–4835, 2009.
[35] C. Bettstetter and C. Hartmann, "Connectivity of wireless multihop networks in a shadow fading environment," *Wireless Networks*, vol. 11, no. 5, pp. 571–579, 2005.
[36] D. Miorandi, E. Alman, and G. Alfano, "The impact of channel randomness on coverage and connectivity of ad hoc and sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 1062–1072, 2008.
[37] X. Ta, G. Mao, and B. D. Anderson, "On the giant component of wireless multi-hop networks in the presence of shadowing," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 5152–1563, 2009.
[38] M. Franceschetti and R. Meester, *Random Networks for Communication: From Statistical Physics to Information Systems*. Cambridge University Press, 2007.

Seh Chun Ng received the BEng degree in IT & Telecommunication from University of Adelaide, Australia, and the MSc degree in IT from Malaysia University of Science and Technology (MUST), Malaysia. He is currently working toward the PhD degree in Engineering at University of Sydney. His research interest include wireless multi-hop networks and graph theory.

Wuxiong Zhang received the BEng degree in Information Security from Shanghai Jiao Tong University (SJTU), Shanghai, P.R. China, in 2008. He is now a PhD candidate of communication and information system in Shanghai Research Center for Wireless Communications, SIMIT, Chinese Academic of Sciences. His research interests lie in the areas of communication theory, information theory and networking. Current research focuses on ad-hoc, and vehicular networks.

Yu Zhang received her Ph.D. degree in computing from Imperial College London, UK, in 2009. She is currently working at Shanghai Engineering Research Center for Broadband Technologies and Applications, China. Her research interests include queueing theory and performance evaluation.

Yang Yang received his PhD degree in Information Engineering from The Chinese University of Hong Kong in 2002. He is currently working at Shanghai Research Center for Wireless Communications (WiCO), SIMIT, Chinese Academy of Sciences. Prior to that, he served the Department of Electronic and Electrical Engineering at University College London (UCL), United Kingdom, as a Senior Lecturer. His general research interests include wireless ad hoc and sensor networks, wireless mesh networks, B3G mobile communication systems.

Guoqiang Mao received PhD in telecommunications engineering in 2002 from Edith Cowan University. He joined the School of Electrical and Information Engineering, the University of Sydney in December 2002 where he is a Senior Lecturer now. His research interests include wireless localization techniques, wireless multihop networks, graph theory and its application in networking, and network performance analysis. He is a Senior Member of IEEE.

Connectivity of Large-Scale CSMA Networks

Tao Yang, *Student member, IEEE,* Guoqiang Mao, *Senior member, IEEE,* and Wei Zhang, *Senior Member, IEEE*

***Abstract*—Wireless multi-hop networks are being increasingly used in military and civilian applications. Connectivity is a prerequisite in wireless multi-hop networks for providing many network functions. In a wireless network with many concurrent transmissions, signals transmitted at the same time will mutually interfere with each other. In this paper we consider the impact of interference on the connectivity of CSMA networks. Specifically, consider a network with n nodes uniformly and *i.i.d.* on a square $\left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$ where a node can only transmit if the sensed power from any other active transmitter is below a threshold, i.e. subject to the carrier-sensing constraint, and the transmission is successful if and only if the SINR is greater than or equal to a predefined threshold. We provide a sufficient condition and a necessary condition, i.e. an upper bound and a lower bound on the transmission power, required for the above network to be asymptotically almost surely (a.a.s.) connected as $n \to \infty$. The two bounds differ by a constant factor only as $n \to \infty$. It is shown that the transmission power only needs to be increased by a constant factor to combat interference and maintain connectivity compared with that considering a unit disk model (UDM) without interference. This result is also in stark contrast with previous results considering the connectivity of ALOHA networks under the SINR model.**

***Index Terms*—Connectivity, CSMA, Wireless Network.**

I. Introduction

Wireless multi-hop networks are being increasingly used in military and civilian applications. Connectivity is a prerequisite in wireless multi-hop networks for providing many network functions (e.g. routing, localization and topology control) [1]–[3]. The scaling behavior of the connectivity property when the network becomes sufficiently large is of particular interest. A wireless multi-hop network is said to be *connected* if and only if (iff) there is at least one (multi-hop) path between any pair of nodes in the network.

Due to the nature of wireless communications, signals transmitted at the same time will mutually interfere with each other. The SINR (signal to interference plus noise ratio) model has been widely used to capture the impact of interference on network connectivity [2], [4], [5]. Under the SINR model, the existence of a directional link between a pair of nodes is determined by the strength of the received signal from the desired transmitter, the interference caused by other concurrent transmissions and the background noise. Assume all nodes use the same transmit power P and let $\boldsymbol{x}_k$, $k \in \Gamma$, be the location of node k, where Γ represents the set of indices of all nodes in the network. A node j can successfully receive the transmitted signal from a node i (i.e. node j *is directly connected to* node i) if the SINR at $\boldsymbol{x}_j$, denoted by $\text{SINR}(\boldsymbol{x}_i \to \boldsymbol{x}_j)$, is above a prescribed threshold β, i.e.

$$\text{SINR}(\boldsymbol{x}_i \to \boldsymbol{x}_j) = \frac{P\ell(\boldsymbol{x}_i, \boldsymbol{x}_j)}{N_0 + \gamma \sum_{k \in \mathcal{T}_i} P\ell(\boldsymbol{x}_k, \boldsymbol{x}_j)} \geq \beta \quad (1)$$

where $\mathcal{T}_i \subseteq \Gamma$ denotes the subset of nodes transmitting at the same time as node i and N_0 is the background noise power. The function $\ell(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is the power attenuation from $\boldsymbol{x}_i$ to $\boldsymbol{x}_j$. The coefficient $0 \leq \gamma \leq 1$ is the inverse of the processing gain of the system and it weighs the impact of interference. In a broadband system using CDMA, γ depends on the orthogonality between codes used during concurrent transmissions and $\gamma < 1$; in a narrow-band system, $\gamma = 1$ [2], [5]. Similarly, node i can receive from node j (i.e. node i is directly connected to node j) iff

$$\text{SINR}(\boldsymbol{x}_j \to \boldsymbol{x}_i) = \frac{P\ell(\boldsymbol{x}_j, \boldsymbol{x}_i)}{N_0 + \gamma \sum_{k \in \mathcal{T}_j} P\ell(\boldsymbol{x}_k, \boldsymbol{x}_i)} \geq \beta. \quad (2)$$

Therefore node i and node j are directly connected, i.e. a bidirectional link exists between node i and node j, iff both (1) and (2) are satisfied.

Dousse *et al.* [5] use the SINR model to analyze the impact of interference on connectivity from the percolation perspective. They consider a network where all nodes are distributed in $\Re^2$ following a homogeneous Poisson point process with a constant intensity λ and an attenuation function ℓ with bounded support. By letting $\mathcal{T}_j = \Gamma / \{i, j\}$, i.e. all other nodes in the network transmit simultaneously with node i irrespective of their relative locations to $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, it is shown that there exists a very small positive constant γ' such that if $\gamma > \gamma'$ there is no infinite connected component in the network, i.e. the network does not *percolate*. Further, when $\gamma < \gamma'$, there exists $0 < \lambda' < \infty$ such that percolation can occur when $\lambda > \lambda'$. An improved result by the same authors in [6] shows that under the more general conditions that $\lambda > \lambda_c$ and the attenuation function has unbounded support, percolation occurs when $\gamma < \gamma'$. Here λ_c is the critical node density above which the network with $\gamma = 0$ (i.e. the unit disk model (UDM) without interference) percolates [7, p48]. These results suggest that percolation under the SINR model can happen iff γ is sufficiently small. They assume that each node transmits randomly and independently, irrespective

T. Yang and G. Mao are with the School of Electrical and Information Engineering, the University of Sydney. Email: {tao.yang, guoqiang.mao}@sydney.edu.au. This research is supported by ARC Discovery project DP110100538, ARC Discovery project DP120102030 and by the Air Force Research Laboratory, under agreement number FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

W. Zhang is with School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, Australia (e-mail: wzhang@ee.unsw.edu.au).

of any nearby transmitter. This corresponds to the ALOHA-type multiple access scheme [2], which however has become obsolete [8].

The more advanced multiple access strategies, e.g. CSMA and CSMA/CD (Carrier Sense Multiple Access with Collision Detection) [9] have become prevailing with widespread adoption. The general idea of CSMA schemes is that nearby nodes will not be scheduled to transmit simultaneously, i.e., a minimum separation distance is imposed among concurrent transmitters**.** Therefore, it is natural to expect that CSMA could improve the performance of ALOHA schemes by alleviating interference, particularly under heavy traffic. On the other hand, this distance constraint leads to a spatial correlation problem which means that the location of a transmitter is dependent on the location of other concurrent transmitters. Therefore, even if all nodes are initially distributed following a Poisson point process (PPP), the set of concurrent transmitters cannot be obtained by independent thinning of the PPP. Thus, the set of concurrent transmitters no longer forms a PPP but a more complicated point process. *Matérn hard-core point process* are widely used to model the set of concurrent transmitters [10]–[12]. However, distribution of such hard-core process is difficult to analyze and a closed-form expression is yet to be obtained [10]–[14]. In this paper, we use an entirely different approach. Particularly by investigating the bounds on interference, instead of an accurate characterization of interference distribution, we are able to avoid the above mentioned difficulty in finding the accurate distribution of concurrent transmitters and the associated interference.

In this paper, we analyze the connectivity of wireless CSMA networks under the SINR model. Specifically, we consider a network with n nodes uniformly *i.i.d.* on a square $\left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$ and each node is capable of performing carrier-sensing operation. A pair of nodes are directly connected iff both (1) and (2) are satisfied. Further, the attenuation function assumes a power-law form, the same model considered in [5], [6]. The contributions of this paper are:

1) We show that the interference experienced by any receiver in the network is upper bounded. Based on this result, we further show that for an arbitrarily chosen SINR threshold, there exists a transmission range R_0 such that a pair of nodes are directly connected if their Euclidean distance is smaller than or equal to R_0. On that basis, we derive a sufficient condition, i.e. an upper bound on the transmission power, for the CSMA network to be *a.a.s.* connected under the SINR model as $n \to \infty$. An event ξ_n depending on n is said to occur *a.a.s.* if its probability tends to 1 as $n \to \infty$.
2) We provide a necessary condition, i.e. a lower bound on the transmission power, for the CSMA network to be *a.a.s.* connected. The lower bound is a tight bound and differs from the upper bound by a constant factor only.
3) We show that the transmission power only needs to be increased by a constant factor to combat interference and maintain connectivity compared with that considering a UDM without interference. This result is in stark contrast with previous results considering the connectivity of ALOHA networks [5], [6] under the SINR model which shows that connectivity is much harder to achieve in the presence of interference and is impossible in a narrow band system where $\gamma = 1$.

The remainder of this paper is organized as follows: Section II reviews related work; Section III defines network and connection models. In Section IV we first derive an upper bound on the interference in CSMA networks. Based on the upper bound, a sufficient condition for connectivity is obtained; Section V investigates a necessary condition for a connected CSMA network; finally Section VI concludes the paper and discusses future work.

II. Related Work

The literature is rich in studying connectivity using the well-known *random geometric graph* and the UDM, which is usually obtained by randomly and uniformly distributing n nodes in a given area and connecting any two nodes iff their Euclidean distance is smaller than or equal to a certain threshold $r(n)$, known as the *transmission range*. This model corresponds to a special case of the SINR model in (1), i.e. when $\gamma = 0$ (perfect orthogonality, no interference). Significant outcomes have been achieved for both asymptotically infinite n [1], [15] and for finite n [16]–[18]. Particularly, Penrose [15] and Gupta and Kumar [1] prove that under the UDM and in a disk of unit area, the above network with a transmission range of $r(n) = \sqrt{\frac{\log n + c(n)}{\pi n}}$ is *a.a.s.* connected as $n \to \infty$ iff $c(n) \to \infty$. Most of the results for finite n are empirical results.

The work [3], [19]–[21] investigate the necessary condition for the above network to be *a.a.s.* connected under the more realistic *log-normal connection model*, where two nodes are directly connected if the received power at one node from the other node, whose attenuation follows the log-normal model [9], is greater than a given threshold. These results however rely on the assumption that the node isolation events are independent, which is yet to be proved.

Despite the significant impact of interference caused by concurrent transmissions on connectivity, limited work exists on analyzing connectivity under the SINR model. In [22], [23], the authors study connectivity from the perspective of channel assignment. Specifically, channel/time slots are assigned to each link for all active links to be simultaneously transmitting while satisfying the SINR requirement. The two papers [5], [6] discussed in Section I study the impact of interference on the connectivity of ALOHA networks from the percolation perspective.

Mao *et al.* [24], [25] study the connectivity problem under a generic *random connection model*, viz. two nodes separated by a Euclidean distance x are directly connected with probability $g(x)$, where $g : [0, \infty) \to [0, 1]$ satisfies the properties of integral boundedness, rotational invariance and non-increasing monotonicity [7], independent of the event that another pair of nodes are directly connected. The authors establish the requirements for an *a.a.s.* connected network.

A major difficulty in moving to the SINR model is that under the unit disk model or the random connection model, connections are assumed to be independent, i.e. the event that a pair of nodes are directly connected and the event that another distinct pair of nodes are directly connected are independent. This independence assumption on connections is critical in the analysis of connectivity under these two models. In the SINR model however, due to the presence of interference, the existence of a direct connection between a pair of nodes depends on both the location and the activities of other nodes in the network.

Some other work exists on modeling the point process formed by concurrent transmitters. The Matérn hard-core point process are widely used to model the set of concurrent transmitters in a CSMA network [10]–[13]. However, such hard-core process are difficult to analyze. Consequently, some work [12], [26], [27] uses PPP to approximate the distribution of concurrent transmitters. A recent work [14] compared the mean interference in hard-core process and in the corresponding PPP approximation.

III. Network Models

We consider a network with n nodes uniformly and *i.i.d.* on a square $\left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$, i.e. the so-called *extended network model* [2], where the network size scales with the network area while the node density is fixed. All nodes use the same transmission power P and there is always a packet at a node waiting to be transmitted. The later assumption allows us to focus on the network property without being disturbed by other factors, e.g. traffic distribution.

A. Attenuation and interference

We consider that the attenuation function ℓ in (1) and (2) only depends on Euclidean distance and is a power-law function [5], [6] $\ell(s) = s^{-\alpha}$ where s represents the Euclidean distance between a pair of nodes and α is the path-loss exponent, which typically varies from 2 to 6 [9, p139]. In this paper we assume $\alpha > 2$. Note that when $\alpha \leq 2$, the interference experienced by a receiver in the CSMA network investigated can not be bounded by a constant. The above assumptions on ℓ are widely used [2], [5], [12] and supported by measurement studies [9]. As commonly done in the connectivity analysis [1], [5]–[7], [15], the impact of small-scale fading is ignored and only bidirectional communication links are considered. Further, since in dense sensor networks and cellular networks the background noise is typically negligibly small [2], [12], we ignore the background noise N_0 in (1) and (2). In addition, we consider that all nodes use the same channel, i.e. $\gamma = 1$, which corresponds to a narrow-band system [2], [5].

B. Carrier-sensing

In CSMA networks, two nodes located at $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ can respectively transmit simultaneously iff they can not detect each other's transmission, i.e. both $P\ell(\boldsymbol{x}_i, \boldsymbol{x}_j)$ and $P\ell(\boldsymbol{x}_j, \boldsymbol{x}_i)$ in (1) and (2) are below a certain detection threshold P_{th}. From the power-law path loss, the carrier-sensing range R_c, which determines the *minimum* Euclidean distance between two concurrent transmitters, is given by

$$R_c = (P/P_{\text{th}})^{1/\alpha} \tag{3}$$

One may alternatively consider a scenario where a node transmits when the aggregated interference is below P_{th}, which forms a trivial extension of the scenario considered in this paper.

IV. A Sufficient Condition for Asymptotically Almost Surely Connectivity

A major challenge in connectivity analysis under the SINR model is that the existence of a direct connection between a pair of nodes depends on both the locations and activities of other nodes in the network, i.e. connections are correlated. In this paper, we resort to a coupling approach to handle the connection correlations. The main idea of coupling technique is to build the connection between a more complicated model and a simpler model with established results such that if a property, e.g. connectivity, is true in the simpler model, it will also be true in the more complicated model. Therefore the property of the more complicated model can be studied by studying the simpler counterpart.

Specifically, we first establish an upper bound on the interference experienced by any receiver in CSMA networks. On that basis, we show that for an arbitrarily chosen SINR threshold, there exists a transmission range R_0 such that a pair of nodes are directly connected if their Euclidean distance is smaller than or equal to R_0. Then we can use existing results on connectivity under the UDM to analyze connectivity under the SINR model.

A. An upper bound on interference and the associated transmission range

The following theorem provides an upper bound on the interference.

Theorem 1. *Consider a CSMA network with nodes distributed arbitrarily on a finite area in $\Re^2$. Denote by r_0 the Euclidean distance between a receiver and its nearest transmitter in the network, which is also the intended transmitter for the receiver. When $r_0 < R_c$, the maximum interference experienced by the receiver is smaller than or equal to $N(r_0) = N_1(r_0) + N_2$, where*

$$N_1(r_0) = \frac{4P\left(\frac{5\sqrt{3}}{4}R_c - r_0\right)^{1-\alpha}\left(\frac{\sqrt{3}}{4}(3\alpha-1)R_c - r_0\right)}{R_c^2(\alpha-1)(\alpha-2)} + \frac{3P}{(R_c - r_0)^{\alpha}} + \frac{3P}{\left(\sqrt{3}R_c - r_0\right)^{\alpha}} + \frac{3P\left(\frac{3}{2}R_c - r_0\right)^{1-\alpha}}{(\alpha-1)R_c} \tag{4}$$

$$N_2 = \frac{3P}{R_c^{\alpha}} + \frac{3P(\frac{3}{2})^{1-\alpha}}{(\alpha-1)R_c^{\alpha}} + \frac{3P}{\left(\sqrt{3}R_c\right)^{\alpha}} + \frac{3P\left(\frac{5}{4}\right)^{1-\alpha}(3\alpha-1)}{(\alpha-1)(\alpha-2)\left(\sqrt{3}R_c\right)^{\alpha}} \tag{5}$$

Proof: See Appendix I. ∎

Remark 2. The upper bound in Theorem 1 is valid for any node distribution. For a sparse network or a network where nodes are placed in a coordinated or planned manner, replacing R_c with the minimum distance among concurrent transmitters, Theorem 1 can be extended to be applicable.

Remark 3. The assumption that $r_0 < R_c$ is valid in most wireless systems which not only require the SINR to be above a threshold and also require the received signal to be of sufficiently good quality. However Theorem 1 does not critically depend on the assumption. For $r_0 \geq R_c$, so long as there exists a positive integer c such that $r_0 < cR_c$ the upper bound can be revised to accommodate the situation by changing the range of the summation in (20) (in Appendix I) from $[3, \infty]$ and $[2, \infty]$ to $[c+2, \infty]$ and $[c+1, \infty]$ respectively and revising the results accordingly.

The following result can be obtained as a ready consequence of Theorem 1.

Corollary 4. *Under the same settings as in Theorem 1, assume that the SINR threshold in* (1) *and* (2) *is* β*. There exists a transmission range* $R_0 < R_c$ *such that a pair of nodes are directly connected if their Euclidean distance is smaller than or equal to* R_0*, given implicitly by*

$$PR_0^{-\alpha}/N(R_0) = \beta \tag{6}$$

Proof: Theorem 1 establishes that the interference experienced by a receiver z at r_0 from its transmitter w, denoted by $I(r_0)$ is upper bounded by $N(r_0)$. Note that, for $r_0 < R_c$, $N(r_0)$ is increasing with r_0 and $Pr_0^{-\alpha}$ is decreasing with r_0. Therefore, using (6) the SINR of a receiver at $r_0 \leq R_0$ from its transmitter, denoted by SINR (r_0), satisfies $\text{SINR}(r_0) = \frac{Pr_0^{-\alpha}}{I(r_0)} \geq \frac{Pr_0^{-\alpha}}{N(r_0)} \geq \beta$.

By symmetry, when the transmission occurs in the opposite direction, i.e. from z to w, the interference generated by the set of nodes that are transmitting at the same time as z is also upper bounded by $N(r_0)$. Therefore the SINR at w is also greater than or equal to β.

Finally the existence of a (unique) solution to (6) can be proved by noting that $\frac{Pr_0^{-\alpha}}{N(r_0)} \to \infty$ as $r_0 \to 0$, $\frac{Pr_0^{-\alpha}}{N(r_0)} \to 0$ as $r_0 \to R_c^-$ and that $\frac{Pr_0^{-\alpha}}{N(r_0)}$ is monotonically decreasing with r_0. ∎

Corollary 4 relates R_0 to P and allows the computation of R_0 given P and the converse. A more convenient way to study the relation between P and R_0 is by noting that $P = P_{\text{th}} R_c^{\alpha}$ and considering R_0 as a function of R_c. Using (4), (5) and letting $\frac{R_c}{R_0} = x$, (6) can be rewritten as

$$\begin{aligned}\frac{1}{\beta} = & \frac{4\left(\frac{5\sqrt{3}}{4}x-1\right)^{1-\alpha}\left(\frac{\sqrt{3}}{4}(3\alpha-1)x-1\right)}{x^2(\alpha-1)(\alpha-2)}\frac{3}{(x-1)^{\alpha}} \\ & + \frac{3}{\left(\sqrt{3}x-1\right)^{\alpha}} + \frac{3\left(\frac{3}{2}x-1\right)^{1-\alpha}}{(\alpha-1)x} + \frac{3(\frac{3}{2})^{1-\alpha}}{x^{\alpha}(\alpha-1)} \\ & + \frac{3}{x^{\alpha}} + \frac{3}{\left(\sqrt{3}x\right)^{\alpha}} + \frac{3\left(\frac{5}{4}\right)^{1-\alpha}(3\alpha-1)}{(\alpha-1)(\alpha-2)\left(\sqrt{3}x\right)^{\alpha}} \end{aligned} \tag{7}$$

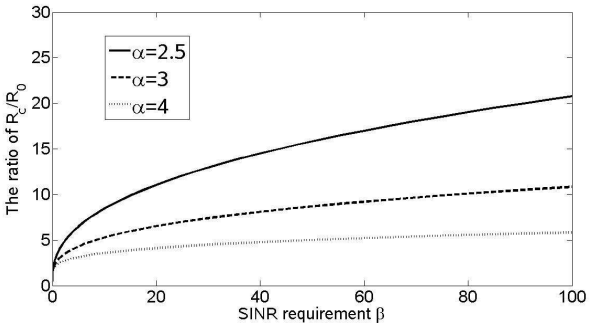


Figure 1. Variation of the ratio $\frac{R_c}{R_0}$ with the SINR requirement β when the path loss exponent α equals to 2.5, 3, 4, respectively.

Figure 1 shows the ratio $\frac{R_c}{R_0}$ as a function of β. Different curves represent different choices of the path loss exponent α. For instance, when $\beta = 10$ and $\alpha = 4$, we have $\frac{R_c}{R_0} = 3.6$.

B. A sufficient condition for connectivity

Based on the transmission range R_0 derived in Corollary 4, we obtain another main result:

Theorem 5. *Consider a CSMA network with a total of* n *nodes* i.i.d. *on a square* $\left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$ *following a uniform distribution. A pair of nodes are directly connected iff both* (1) *and* (2) *(*$\gamma = 1$ *and* $N_0 = 0$ *in* (1) *and* (2) *) are satisfied. As* $n \to \infty$*, the above network is* a.a.s. *connected if the transmission power*

$$P = P_{th} b_1^{\alpha} \left(\log n + c(n)\right)^{\frac{\alpha}{2}}, \tag{8}$$

where $b_1 = b'/\sqrt{\pi}$*,* $c(n) = o(\log n)$ *and* $c(n) \to \infty$ *as* $n \to \infty$ *and* $\infty > b' > 1$ *is the solution to* (7) *(By* $f(x) = o(g(x))$*, we mean that* $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$*.).*

Proof: The results in [1], [15] show that, for a network with a total of n nodes uniformly *i.i.d.* on a $\sqrt{n} \times \sqrt{n}$ square and a pair of nodes are directly connected iff their Euclidean distance is smaller than or equal to a given threshold $r(n)$ (i.e., UDM), the network is *a.a.s.* connected as $n \to \infty$ iff $r(n) = \sqrt{\frac{\log n + c(n)}{\pi}}$ where $c(n) \to \infty$ as $n \to \infty$. Using this result, (7) (letting $b' = \frac{R_c}{R_0}$), Corollary 4 and Theorem 1, the result in the theorem follows. ∎

The implication of Theorem 5 is that in CSMA networks, since the interference is bounded above by a constant almost surely as shown in Theorem 1, to meet an arbitrarily high (albeit constant with the increase in n) β, the power only needs to be increased by a constant factor compared with that in the unit disk model to maintain the same set of connections. This result is in contrast to the ALOHA networks considered in [5], [6] in which percolation only occurs for a sufficiently small γ.

V. A Necessary Condition for Asymptotically Almost Surely Connectivity

Section IV derives a sufficient condition for a connected CSMA network as $n \to \infty$ in the presence of interference.

A logical question arises: what is the necessary condition for the same CSMA network to be connected as $n \to \infty$.

In a CSMA network, any set of nodes can transmit simultaneously as long as the carrier-sensing constraints are satisfied. Further, in a large-scale network, scheduling is often performed in a distributed manner. In the absence of accurate global knowledge on which particular set of nodes are simultaneously transmitting at a particular time instant, it is natural that a node sets its transmission power to be above *the minimum transmission power* required for a network to be connected under *any* scheduling algorithm (It is trivial to show that, see also the proof of Lemma 6, when the transmission power increases, connectivity will also improve). Denote that minimum power by P'_Ω where Ω represents the set of all scheduling algorithms satisfying the carrier-sensing constraints. In this section, we investigate P'_Ω, i.e. a necessary condition required for connectivity as $n \to \infty$. This is done by analyzing the transmission power required for the above network to have no isolated node which is a necessary condition for having a connected network. The following lemma is required for the analysis of P'_Ω:

Lemma 6. *Denote by P_Ω (respectively, P_ω) the minimum transmission power required for the network to have no isolated node under any scheduling (respectively, under a particular scheduling ω). We have $P'_\Omega \geq P_\Omega = \max_{\omega\in\Omega} P_\omega$.*

Proof: We prove the lemma by showing that the minimum transmission power required for the network to have no isolated node under *any* scheduling has to be greater than or equal to the minimum transmission power required for the same network to have no isolated node under a particular scheduling.

Define a set of nodes that can simultaneously transmit while satisfying the carrier-sensing constraints as an *independent set*. Obviously the independent set depends on the transmission power of nodes. As the transmission power decreases, other things being equal, R_c will decrease and the number of nodes that can simultaneously transmit will increase or remain the same.

Denote by ϕ' a set of nodes that are scheduled to transmit simultaneously in the CSMA network. It follows that ϕ' must be an independent set. *Given* ϕ', a node $v \in \phi'$ is *isolated* if there is no node in the network that can successfully receive from it when the nodes in ϕ' are simultaneously transmitting. Further, as explained in the last paragraph, the independent set depends on the transmission power. When the transmission power is decreased from P_1 to P_2, where $P_2 \leq P_1$, if ϕ' is an independent set at power level P_1, it will also be an independent set at power level P_2. Based on the above observation and using (1) and (2), a decrease in the transmission power will cause a decrease in the SINR, it readily follows that if a node $v \in \phi'$ is isolated at power level P_1 when the set of active transmitters is ϕ', it will also be isolated at power level P_2 when the set of active transmitters is ϕ'. For any transmission power less than $P_\Omega = \max_{\omega\in\Omega} P_\omega$, there exists a scheduling that will result the network to have an isolated node at that power level. Therefore, P_Ω has to satisfy $P_\Omega = \max_{\omega\in\Omega} P_\omega$. ■

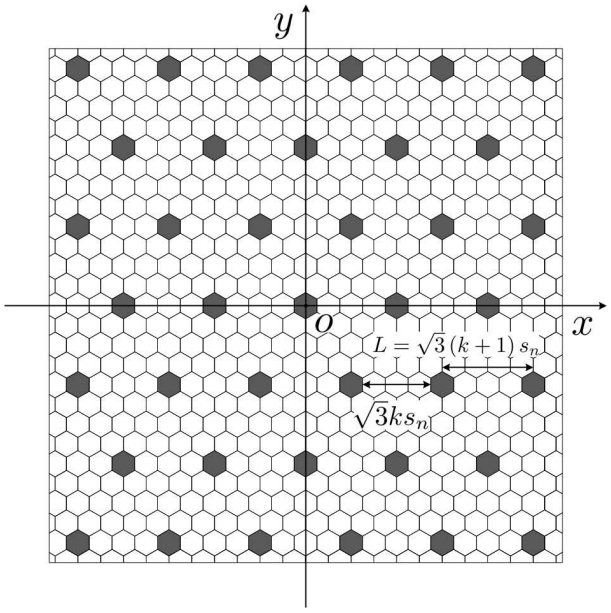


Figure 2. An illustration of the hexagonal partition of the network area. The shaded hexagons represent simultaneously active hexagons, where $k = 3$.

Remark 7. As an easy consequence of Lemma 6, the probability that a CSMA network has no isolated node is a non-increasing function of the transmission power.

Now the task becomes constructing a particular scheduling which gives as large P_ω as possible, i.e. a tight lower bound on P'_Ω. Next we construct such scheduling ω heuristically.

A. Construction of scheduling algorithm ω

Obviously, ω needs to satisfy the constraint on the minimum separation distance between concurrent transmitters imposed by the carrier-sensing requirement. Meanwhile, ω needs to schedule as many concurrent transmissions as possible to maximize interference, hence P_ω.

We start with a lemma that is required for the construction of ω.

Lemma 8. *Partition the square $\left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$ into non-overlapping hexagons of equal side length s_n such that the origin o coincides with the centre of a hexagon and two diagonal vertices of this hexagon, whose Euclidean distance is $2s_n$, are located on y axis, as shown in Figure 2. We call a hexagon an* interior hexagon *if it is entirely contained in the square $\left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$. When $s_n = \sqrt{(2\log n)/5}$, a.a.s. each interior hexagon is occupied by at least one node as $n \to \infty$.*

Proof: Because nodes are uniformly *i.i.d.*, the probability that an arbitrary interior hexagon is empty is $\left(1 - \frac{3\sqrt{3}s_n^2}{2n}\right)^n$. Let ξ_i be the event that an interior hexagon i is empty, where $i \in \Xi$ and Ξ denotes the set of indices of all interior hexagons. There are at most $\frac{2n}{3\sqrt{3}s_n^2}$ interior hexagons.

Denote by A_n the event that there is at least one empty interior hexagon in $\left[-\frac{\sqrt{n}}{2},\frac{\sqrt{n}}{2}\right]^2$. It follows that $\Pr(A_n) = \Pr\left(\cup_{i\in\Xi}\xi_i\right)$. Using union bound, we have $\Pr\left(\cup_{i\in\Xi}\xi_i\right) \leq \sum_{i\in\Xi}\Pr(\xi_i) \leq \frac{2n\left(1-\frac{3\sqrt{3}s_n^2}{2n}\right)^n}{3\sqrt{3}s_n^2}$. Using the fact that $1-x \leq \exp(-x)$ and $s_n = \sqrt{\frac{2\log n}{5}}$, we have $\lim_{n\to\infty}\Pr(A_n) \leq \lim_{n\to\infty}\frac{2ne^{-\frac{3\sqrt{3}s_n^2}{2}}}{3\sqrt{3}s_n^2} = \lim_{n\to\infty}\frac{5n}{3\sqrt{3}n^{\frac{3\sqrt{3}}{5}}\log n} = 0$ which completes the proof. ∎

Hereinafter, we declare a hexagon to be *active* if there is a node transmitting in it. We consider a scheduling ω that uses the hexagons as the basic unit for scheduling. Due to the minimum separation distance constraint, any two simultaneously active hexagons should be separated by a minimum Euclidean distance (depending on the carrier-sensing range given in (3)). Let k be an integer and represent the minimum number of inactive hexagons between two closest simultaneously active hexagons (see Figure 2). Any two nodes inside the two active hexagons are separated by a Euclidean distance of at least $\sqrt{3}ks_n$. With a bit twist of terminology, we further define a *maximal independent set* for scheduling to be the set of hexagons that a) includes as many hexagons as possible; and b) closest hexagons in the set are separated by exactly k adjacent hexagons. Figure 2 illustrates such a maximal independent set with $k=3$.

We define ω such that only hexagons belonging to the same maximal independent set can be active at the same time. No nodes in the same hexagon can be scheduled to transmit simultaneously. (Note that if a hexagon intersecting the border of $\left[-\frac{\sqrt{n}}{2},\frac{\sqrt{n}}{2}\right]^2$ has node(s) in it, it is also included into the maximal independent set and its node(s) are treated in the same way as other nodes in interior hexagons.) As a consequence of the CSMA constraint and the definition of k,

$$\sqrt{3}ks_n \geq R_c \geq \sqrt{3}(k-1)s_n \tag{9}$$

B. Probability of having no isolated node

In this subsection, we derive a lower bound on P_ω for ω defined in the previous subsection. This is done by analyzing the event that the network has no isolated node under ω. The following theorem summarizes another major outcome of the paper:

Theorem 9. *Under the same setting in Theorem 5 and the scheduling algorithm ω, a necessary condition on P_ω for the CSMA network to have no isolated node a.a.s. as $n\to\infty$ is*

$$P_\omega \geq P_{th}b_2^\alpha(\log n)^{\frac{\alpha}{2}} \tag{10}$$

where $b_2 = \sqrt{6/5}(b-1)$ and b is the smallest integer satisfying the inequality: $\frac{2\left(\sqrt{3}(b+1)+1\right)^{1-\alpha}\left(\sqrt{3}(\alpha-1)(b+1)+1\right)}{(b+1)^2(\alpha-1)(\alpha-2)} \leq \frac{1}{\beta}\left(\frac{2\pi}{5}\right)^{\frac{\alpha}{2}}$.

Proof: The main strategy used is to couple the network under the SINR model with the associated network under UDM. Then, an upper bound on the probability of having no isolated node in the network under the SINR model is obtained by using existing results for UDM.

Denote the Euclidean distance between the centers of two closest hexagons in a maximal independent set by $L = \sqrt{3}(k+1)s_n$. See Figure 2 for an illustration. Divide the hexagons belonging to the same maximal independent set as a hexagon h_i into tiers of increasing Euclidean distance from the centre of h_i using a similar strategy as that in the proof of Theorem 1. The m^{th} tier of h_i has at most $6m$ hexagons. Further, we declare that the m^{th} tier of h_i is *complete in a given area* if all the $6m$ hexagons are entirely enclosed in this given area. Denote by C_A a square $\left[-\frac{\sqrt{cn}}{2},\frac{\sqrt{cn}}{2}\right]^2$ ($0<c<1$ and the exact value of c will be decided later in this paragraph). The hexagon containing the origin o has a number of $t = \left\lfloor\frac{\frac{c\sqrt{n}}{2}-\frac{\sqrt{3}s_n}{2}}{L}\right\rfloor$ complete tiers in C_A. As c increases, t increases as well. For the hexagons located in C_A but near the border of C_A, the number of complete tiers in the square $\left[-\frac{\sqrt{n}}{2},\frac{\sqrt{n}}{2}\right]^2$ decreases with an increase in c. We choose the value of c such that each hexagon inside C_A has at least t complete tiers in the square $\left[-\frac{\sqrt{n}}{2},\frac{\sqrt{n}}{2}\right]^2$, and the value of t is maximized. Let C_A' be the union of hexagons entirely contained in C_A. With a little bit abuse of terminology, we use C_A (C_A') to denote both the area itself and the size of the area. We can obtain $\lim_{n\to\infty}\frac{C_A'}{C_A}=1$.

Consider an arbitrarily node i transmitting inside a hexagon h_i in C_A'. If there is no node that can receive from it, then node i is isolated. Let I_{min} be the *minimum* interference that could possibly be experienced by a potential receiver of node i under ω. Note that the Euclidean distance between the transmitter inside a hexagon in the m^{th} tier of h_i and the centre of hexagon h_i is less than $mL+s_n$ (see Figure 2). Using Lemma 12 gives

$$\begin{aligned}
I_{min} &\geq \sum_{m=1}^{t}6m(mL+s_n)^{-\alpha}P \\
&= 6Ps_n^{-\alpha}\sum_{m=1}^{t}m\left(\sqrt{3}m(k+1)+1\right)^{-\alpha} \\
&= 6Ps_n^{-\alpha}\int_1^t\lfloor x\rfloor\left(\sqrt{3}\lfloor x\rfloor(k+1)+1\right)^{-\alpha}dx \quad (11)\\
&\geq 6Ps_n^{-\alpha}\int_1^t x\left(\sqrt{3}x(k+1)+1\right)^{-\alpha}dx \quad (12)
\end{aligned}$$

where $\lfloor x\rfloor$ denotes the largest integer smaller than or equal to x. (12) is obtained due to the fact that $x\left(\sqrt{3}x(k+1)+1\right)^{-\alpha}$ is a decreasing function when $x > \frac{1}{\sqrt{3}(k+1)(\alpha-1)}$ and $\sqrt{3}(k+1)(\alpha-1) > 1$ for $\alpha>2$ and $k\geq 1$. Therefore $x\left(\sqrt{3}x(k+1)+1\right)^{-\alpha}$ is a decreasing function when $x>1$. Further, noting that $\lim_{n\to\infty}t = \lim_{n\to\infty}\left\lfloor\frac{\frac{\sqrt{cn}}{2}-\frac{\sqrt{3}s_n}{2}}{L}\right\rfloor = \infty$, it follows that

$$\begin{aligned}
&\lim_{n\to\infty}6\int_1^t x\left(\sqrt{3}x(k+1)+1\right)^{-\alpha}dx \\
=&\frac{2\left(\sqrt{3}(k+1)+1\right)^{1-\alpha}\left(\sqrt{3}(\alpha-1)(k+1)+1\right)}{(k+1)^2(\alpha-1)(\alpha-2)} \\
\triangleq& f(k)
\end{aligned}$$

The above equation implies that for an arbitrarily small positive constant ε, there exists a positive integer n_ε such that when $n \geq n_\varepsilon$

$$\text{RHS of (12)} \quad \geq \quad P s_n^{-\alpha} \left(f\left(k\right) - \varepsilon \right) \triangleq J_n \qquad (13)$$

Let d be the Euclidean distance between node i and its receiver. By (1), (2), it follows that *only* when $\frac{P d^{-\alpha}}{J_n} \geq \beta$, the transmission from node i to its receiver could *possibly* be successful. In other words, if there is no node within a Euclidean distance of $R = \left(\beta J_n / P\right)^{-\frac{1}{\alpha}}$ to node i, then it is isolated.

Denote by M and M^{SINR} the (random) number of isolated nodes in the CSMA network in the square $\left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$ and in $C_A' \subset \left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$ respectively. Denote by M^{UDM} the (random) number of isolated nodes in an area $C_A' \subset \left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$ in a network with a total of n nodes uniformly i.i.d. on the square $\left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$ under UDM with the transmission range R. Based on the discussion in the last paragraph and using the coupling technique [7], it can be shown that $\Pr\left(M \geq 1\right) \geq \Pr\left(M^{\text{SINR}} \geq 1\right) \geq \Pr\left(M^{\text{UDM}} \geq 1\right)$. Consequently,

$$\Pr\left(M = 0\right) \leq \Pr\left(M^{\text{UDM}} = 0\right) \qquad (14)$$

It remains to find the value of $\Pr\left(M^{\text{UDM}} = 0\right)$. We first consider a network with a total of n nodes distributed on a square $\left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$ under UDM with a transmission range $r\left(n\right)$. It is well-known that when the average node degree in the above network equals to $\log n + \zeta\left(n\right)$ and $\lim_{n\to\infty} \zeta\left(n\right) = \zeta$ where ζ is a constant ($\zeta = \infty$ is allowed), the probability that there is no isolated node in the above network asymptotically converges to $e^{-e^{-\zeta}}$ as $n \to \infty$ [7], [28], [29]. Further, it was shown in [30] that boundary effect has an asymptotically vanishingly impact on the number of isolated nodes. Let Z be a random integer representing the number of nodes located inside $C_A \subset \left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right]^2$. $\mathrm{E}\left(Z\right) = cn$ and $\mathrm{Var}\left(Z\right) = cn\left(1-c\right)$. Let $M^{r(n)}$ be the number of isolated nodes within an area C_A in the above network with a transmission range $r\left(n\right)$. Based on the above results, conditioned on that $Z = cn$ we have (here we have omitted some trivial discussions involving the situation that cn is not an integer)

$$\lim_{n\to\infty} \Pr\left(M^{r(n)} = 0 \middle| \, Z = cn \right) = e^{-ce^{-\zeta}} \qquad (15)$$

Using Chebyshev's inequality, for $0 < \delta < \frac{1}{2}$, we obtain that

$$\lim_{n\to\infty} \Pr\left(\left| Z - cn \right| \geq \left(cn\right)^{\frac{1}{2}+\delta} \right) \leq \lim_{n\to\infty} \frac{\mathrm{Var}\left(Z\right)}{\left(\left(cn\right)^{\frac{1}{2}+\delta}\right)^2} = 0 \qquad (16)$$

Let $f\left(n\right) = \left(cn\right)^{\frac{1}{2}+\delta}$. Using the following two equations: $\log\left(n + f\left(n\right)\right) + \zeta\left(n\right) = \log n + \log\left(1 + \frac{f(n)}{n}\right) + \zeta\left(n\right)$ and $\lim_{n\to\infty} \log\left(1 + \frac{f(n)}{n}\right) + \zeta\left(n\right) = \lim_{n\to\infty} \zeta\left(n\right) = \zeta$ and (15), it can be shown that $\lim_{n\to\infty} \Pr\left(M^{r(n)} = 0 \middle| \, Z = cn + f\left(n\right) \right) = e^{-ce^{-\zeta}}$. Hence, for any integer m satisfying $-f\left(n\right) \leq m \leq f\left(n\right)$, $\lim_{n\to\infty} \Pr\left(M^{r(n)} = 0 \middle| \, Z = cn + m \right) = e^{-ce^{-\zeta}}$. This equation, together with (16), allows us to conclude that when $r\left(n\right) = \sqrt{\frac{\log n + \zeta(n)}{\pi}}$

$$\lim_{n\to\infty} \Pr\left(M^{r(n)} = 0 \right) = e^{-ce^{-\zeta}} \qquad (17)$$

As a result of (14), a necessary condition for $\lim_{n\to\infty} \Pr\left(M = 0\right) = 1$ is that $\lim_{n\to\infty} \Pr\left(M^{\text{UDM}} = 0\right) = 1$. Using the fact that $\lim_{n\to\infty} \frac{C_A'}{C_A} = 1$ and (17), it follows that a necessary condition for the network under the SINR model to *a.a.s.* have no isolated node is that $R \geq \sqrt{\frac{\log n + \zeta(n)}{\pi}}$ and $\zeta\left(n\right) \to \infty$ as $n \to \infty$. As denoted $R = \left(\beta J_n / P\right)^{-\frac{1}{\alpha}}$, together with the value of J_n in (13) and the value of s_n in Lemma 8 , we obtain that $f\left(k\right) \leq \frac{1}{\beta}\left(\frac{2\pi}{5}\frac{\log n}{\log n + \zeta(n)}\right)^{\frac{\alpha}{2}} + \varepsilon$. Letting $n \to \infty$ and then $\varepsilon \to 0$ in the above inequality yields $f\left(k\right) \leq \frac{1}{\beta}\left(\frac{2\pi}{5}\right)^{\frac{\alpha}{2}}$. Based on the above equation, together with (3) and (9), Theorem 9 results. ∎

The following corollary is obtained as a ready consequence of Theorem 9 and Lemma (6).

Corollary 10. *A necessary condition required for CSMA networks to be a.a.s. connected as $n \to \infty$ under* any *scheduling algorithm, i.e. a lower bound on P_Ω', is given by*

$$P_\Omega' \geq P_{\text{th}} b_2^\alpha \left(\log n\right)^{\frac{\alpha}{2}} \qquad (18)$$

Comparing the lower bound on P_Ω' in (18) with the upper bound in (8) and noting that $c\left(n\right) = o\left(\log n\right)$, it can be shown that, given an arbitrary β, the two bounds differ by a constant factor only as $n \to \infty$. Figure 3 shows the plot of the two constant factors, viz. b_1 and b_2, in (8) and in (18) respectively as a function of β when $\alpha = 4$. The curve representing b_2 is a step function due to the granularity caused by the integer k in the scheduling algorithm ω.

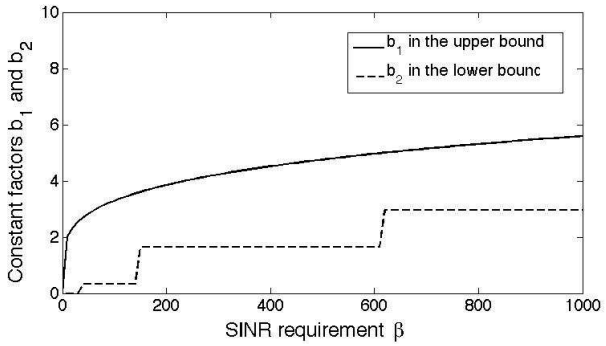


Figure 3. A plot of the two constant factors b_1 and b_2 in the upper bound (8) and in the lower bound (18) when $\alpha = 4$.

VI. Conclusion and Future Work

In this paper, we studied the connectivity of wireless CSMA networks considering the impact of interference. We showed

that, different from an ALOHA network, the aggregate interference experienced by any receiver in CSMA networks is upper bounded even when the coefficient γ in (1) and (2) equals to 1.

An upper bound and a lower bound were obtained on the critical transmission power required for having an *a.a.s.* connected CSMA network. The two bounds are tight and differ by a constant factor only. The results suggested that any pair of nodes can be connected for an arbitrarily high SINR requirement so long as the carrier-sensing capability is available. Compared with that considering UDM without interference, the transmission power only needs to be increased by a constant factor to combat interference and maintain connectivity. This is a optimistic result compared with previous results on the connectivity of ALOHA networks under the SINR model.

The gap between the two bounds can be further narrowed by considering more complicated geometric shapes than hexagons. However such improvement is possibly of minor importance. The implication of the results in this paper is that there exists a spatial and temporal scheduling algorithm in a large scale CSMA network that allows as many as possible concurrent transmissions, and meanwhile, allows any pair of nodes in the network to be connected under an arbitrarily high SINR requirement. We also introduce a hexagon-based scheduling algorithm that allows the CSMA network to be connected. However, it remains a major challenge to find the optimum scheduling algorithm that gives the minimum delay and the maximum capacity under a specific traffic distribution.

APPENDIX I PROOF OF THEOREM 1

A network on a finite area, denoted by $A \subset \Re^2$, can always be obtained from a network on an infinite area $\Re^2$ with the same node density and distribution by removing these nodes outside A. Such removal process will also remove all transmitters outside A. Therefore the interference at a receiver in A is less than or equal to the interference experienced by its counterpart in a network in $\Re^2$. It then suffices to show that the interference in a network in $\Re^2$ is bounded.

Consider that an arbitrary receiver z is located at a Euclidean distance r_0 from its closest transmitter w, which is also the intended transmitter for z. We construct a coordinate system such that the origin of the coordinate system is at w and z is on the $+y$ axis, as shown in Fig. 4.

In a CSMA network, the distance between any two concurrent transmitters is at least R_c. Draw a circle of radius $R_c/2$ centered at each transmitter. Then the two circles centered at two closest transmitters cannot overlap except at a single point. Therefore the problem of determining the maximum interference can be transformed into one that determining the maximum number of equal-radius non-overlapping circles that can be packed into $\Re^2$. The densest circle packing, i.e. fitting the maximum number of non-overlapping circles into $\Re^2$, is obtained by placing the circle centers at the vertices of a hexagonal lattice [31, p. 8], as shown in Fig. 4.

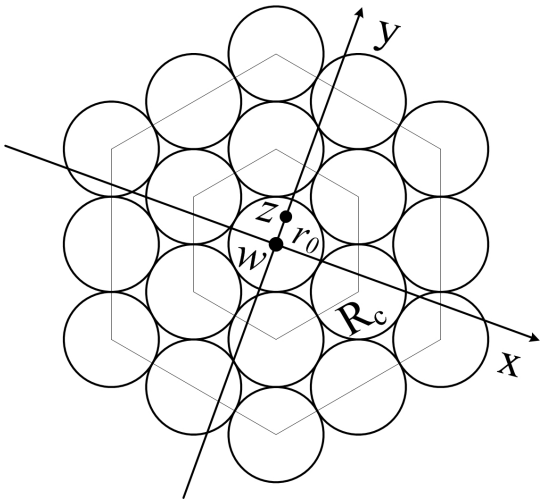


Figure 4. An illustration of the densest equal-circle packing.

Group the vertices of the hexagonal lattice into tiers of increasing distances from the origin. The six vertices of the first tier are within a Euclidean distance R_c to the origin. The $6m$ vertices in the m^{th} tier are located at distances within $((m-1)R_c, mR_c]$ from the origin.

Let I_1 be the interference caused by transmitters, hereinafter referred to as interferers in this section, above the x-axis at node z. Using the triangle inequalities gives $\|\boldsymbol{x}_i - \boldsymbol{z}\| \geq \|\boldsymbol{x}_i\| - r_0$ where $\boldsymbol{x}_i$ is the location of an interferer above the x-axis. Among the $6m$ interferers in the m^{th} group, half of them are located above the x-axis. Among these interferers in the m^{th} group above the x-axis, three of them are at a Euclidean distance of exactly mR_c from the origin and the rest $3(m-1)$ interferers are at Euclidean distances within $[\frac{\sqrt{3}}{2}mR_c,\ mR_c]$. Hence, we have

$$I_1 \leq \sum_{m=1}^{\infty}\left(\frac{3(m-1)P}{\left(\frac{\sqrt{3}}{2}mR_c - r_0\right)^{\alpha}} + \frac{3P}{(mR_c - r_0)^{\alpha}}\right) \tag{19}$$

Look at the first summation in (19). Let $U_m, m = 3, \ldots, \infty$, be random variables uniformly and i.i.d. in $[m-1/2,\ m+1/2]$. It follows from the convexity of $\frac{3(m-1)P}{\left(\frac{\sqrt{3}}{2}mR_c - r_0\right)^{\alpha}}$ and Jensen's inequality (used in the second step) that

$$\begin{aligned}
&\sum_{m=3}^{\infty}\frac{3(m-1)P}{\left(\frac{\sqrt{3}}{2}mR_c - r_0\right)^{\alpha}} \\
=\ &\sum_{m=3}^{\infty}\frac{3(\mathrm{E}(U_m)-1)P}{\left(\frac{\sqrt{3}}{2}\mathrm{E}(U_m)R_c - r_0\right)^{\alpha}} \\
\leq\ &\sum_{m=3}^{\infty}\mathrm{E}\left(\frac{3(U_m-1)P}{\left(\frac{\sqrt{3}}{2}U_mR_c - r_0\right)^{\alpha}}\right) \\
=\ &\sum_{m=3}^{\infty}\int_{m-1/2}^{m+1/2}\frac{3(x-1)P}{\left(\frac{\sqrt{3}}{2}xR_c - r_0\right)^{\alpha}}dx \\
=\ &3P\int_{5/2}^{\infty}(x-1)\left(\frac{\sqrt{3}}{2}xR_c - r_0\right)^{-\alpha}dx
\end{aligned} \tag{20}$$

$$= \frac{4P\left(\frac{5\sqrt{3}}{4}R_c - r_0\right)^{1-\alpha}\left(\frac{\sqrt{3}}{4}(3\alpha-1)R_c - r_0\right)}{R_c^2(\alpha-1)(\alpha-2)} \quad (21)$$

Likewise, we also have $\sum_{m=2}^{\infty}\frac{3P}{(mR_c-r_0)^{\alpha}} \leq \frac{3P\left(\frac{3}{2}R_c-r_0\right)^{1-\alpha}}{(\alpha-1)R_c}$. As a result of the last equation and (19), (21), (4), it follows that $I_1 \leq N_1(r_0)$.

Now we consider the total interference caused by interferers below the x-axis at node z, denoted by I_2. Let $\boldsymbol{x}_i$ be the location of an interferer below the x-axis, it follows from the triangle inequality that $\|\boldsymbol{x}_i - \boldsymbol{z}\| \geq \|\boldsymbol{x}_i\|$. Therefore

$$\begin{aligned} I_2 &\leq \sum_{m=1}^{\infty}\left(\frac{3P}{(mR_c)^{\alpha}} + \frac{3(m-1)P}{\left(\frac{\sqrt{3}}{2}mR_c\right)^{\alpha}}\right) \\ &\leq \frac{3P}{R_c^{\alpha}} + \frac{3P(\frac{3}{2})^{1-\alpha}}{(\alpha-1)R_c^{\alpha}} + \frac{3P}{\left(\sqrt{3}R_c\right)^{\alpha}} \\ &\leq +\frac{3P\left(\frac{5}{4}\right)^{1-\alpha}(3\alpha-1)}{(\alpha-1)(\alpha-2)\left(\sqrt{3}R_c\right)^{\alpha}} \end{aligned} \quad (22)$$

Combining $I_1 \leq N_1(r_0)$ and (22), Theorem 1 is proved.

APPENDIX II LEMMA 12

Lemma 12 is needed in the proof of Theorem 9. Theorem 11 is used to prove Lemma 12.

Theorem 11. *(Theorem 1 in [32]) Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_j$ be j arbitrary points in $\Re^2$. Let $w_1, w_2, \ldots w_j$ be j positive numbers regarded as weights attached to these points, and define a position vector $\boldsymbol{c}$ by $\sum_{i=1}^{j} w_i \boldsymbol{v}_i = W\boldsymbol{c}$ where $W = \sum_{i=1}^{j} w_i$. Then for an arbitrary point $\boldsymbol{z}$, the following holds: $\sum_{i=1}^{j} w_i \|\boldsymbol{v}_i - \boldsymbol{z}\|^2 = \sum_{i=1}^{j} w_i \|\boldsymbol{v}_i - \boldsymbol{c}\|^2 + W\|\boldsymbol{z} - \boldsymbol{c}\|^2$*

Lemma 12. *Consider a triangular lattice with unit side length and having a vertex located at the origin o. Define the 1^{st} tier of points to be the six points placed at the vertices of the triangular lattice at a distance of 1 to the origin o. Let the m^{th} tier of points be the $6m$ points placed at the vertices of the triangular lattice located at distances within $(m-1,\ m]$ from the origin o, as shown in Figure 5. The total number of points from the 1^{st} tier to the m^{th} tier then equals to $j = 3m(1+m)$. Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots \boldsymbol{v}_j$ be the location vectors of these j points and the points are ordered according to their distances to the origin o in a non-decreasing order. For an arbitrary point $\boldsymbol{z}$ located inside the hexagon formed by the 1^{st} tier of six points, the following holds: $\sum_{i=1}^{j}\|\boldsymbol{v}_i - \boldsymbol{z}\|^{-\alpha}$ is minimized when $\boldsymbol{z}$ is located at the origin o.*

Proof: Now we use Theorem 11 to prove Lemma 12. Letting all attached weights w_i equal to 1 and using Theorem 11, for an arbitrary point $\boldsymbol{z}$ located inside the hexagon formed by the 1st tier of six points, we have

$$\sum_{i=1}^{6}\|\boldsymbol{v}_i - \boldsymbol{z}\|^2 = \sum_{i=1}^{6}\|\boldsymbol{v}_i - \boldsymbol{c}\|^2 + 6\|\boldsymbol{z} - \boldsymbol{c}\|^2 \quad (23)$$

where $\boldsymbol{c}$ is given by $\sum_{i=1}^{6}\boldsymbol{v}_i = 6\boldsymbol{c}$. It is clear that $\boldsymbol{c}$ is the centroid of the six points. Since the hexagon has a unit side length, $\|\boldsymbol{v}_i - \boldsymbol{c}\|$ equals to 1. Let $x_i = \|\boldsymbol{v}_i - \boldsymbol{z}\|$ and $y = \|\boldsymbol{z} - \boldsymbol{c}\|$. The problem in Lemma 12 can then be converted to the following constrained minimization problem:

$$\begin{aligned} \text{minimize} \quad & f(x_1, \ldots, x_6) = \textstyle\sum_{i=1}^{6} x_i^{-\alpha} \\ \text{subject to} \quad & h(x_1, \ldots, x_6) = \textstyle\sum_{i=1}^{6} x_i^2 - 6 - 6y^2 = 0 \end{aligned}$$

where the constraint is due to (23). Using the method of Lagrange multipliers, we first construct the Lagrangian in the following: $F(x_1, \ldots x_6, \Lambda) = f(x_1, \ldots, x_6) + \Lambda h(x_1, \ldots, x_6)$ where the parameter Λ is known as the Lagrange multiplier. Then find the gradient and set it to zero: $\nabla F(x_1, \ldots x_6, \Lambda) = \begin{pmatrix} -\alpha x_1^{-\alpha-1} + 2\Lambda x_1 \\ \vdots \\ -\alpha x_6^{-\alpha-1} + 2\Lambda x_6 \\ h(x_1, x_2, \ldots, x_6) \end{pmatrix}^T = \mathbf{0}$. Solving the above equation, it is obtained that $\Lambda = \frac{\alpha}{2}\left(1+y^2\right)^{-\frac{\alpha+2}{2}}$ and $x_1 = x_2 \ldots = x_6 = \left(\frac{2\Lambda}{\alpha}\right)^{\frac{-1}{\alpha+2}} = \left(1+y^2\right)^{\frac{1}{2}}$. Since $x_i = \|\boldsymbol{v}_i - \boldsymbol{z}\|$ denotes the Euclidean distance from $\boldsymbol{v}_i$ to $\boldsymbol{z}$, only when $\boldsymbol{z} = \boldsymbol{c}$, we can have $x_1 = x_2 = \ldots = x_6 = 1$. It follows that the minimum of $f(x_1, x_2, \ldots, x_6)$ is obtained only when $\boldsymbol{z}$ is located at the origin o. Further, for the $6m$ points of the m^{th} tier, using the same method, it can be shown that $\sum_{i=1}^{6m}\|\boldsymbol{v}_i - \boldsymbol{z}\|^{-\alpha}$ is minimized only when $\boldsymbol{z}$ is located at the origin o. The result follows. ■

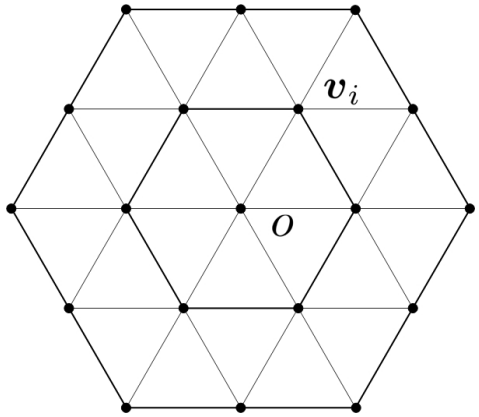


Figure 5. An illustration of a triangular lattice

REFERENCES

[1] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity," in *Proc. IEEE Conference on Decision and Control*, 1998, pp. 1106–1110.

[2] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, 2009.

[3] C. Bettstetter and C. Hartmann, "Connectivity of wireless multihop networks in a shadow fading environment," *Wireless Networks*, vol. 11, no. 5, pp. 571–579, 2005.

[4] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[5] O. Dousse, F. Baccelli, and P. Thiran, "Impact of interferences on connectivity in ad hoc networks," *IEEE/ACM Transactions on Networking*, vol. 13, no. 2, pp. 425–436, 2005.

[6] O. Dousse, M. Franceschetti, N. Macris, R. Meester, and P. Thiran, "Percolation in the signal to interference ratio graph," *Journal of Applied Probability*, vol. 43, no. 2, pp. 552–562, 2006.

[7] M. Franceschetti and R. Meester, *Random Networks for Communication from Statistical Physics to Information Systems*. Cambridge University Press, 2007.

[8] S. Kumar, V. S. Raghavan, and J. Deng, "Medium access control protocols for ad hoc wireless networks: A survey," *Ad Hoc Networks*, vol. 4, no. 3, pp. 326–358, 2006.

[9] T. S. Rappaport, *Wireless Communications Principles and Practice*, 2nd ed. Prentice Hall, 2002.

[10] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks Volume II : APPLICATION*. Paris: Now Publishers, 2009.

[11] A. Busson, G. Chelius, and Acm, "Point processes for interference modeling in csma/ca ad-hoc networks," in *Pe-Wasun09: Proceedings of the Sixth Acm International Symposium on Performance Evaluation of Wireless Ad-Hoc, Sensor, and Ubiquitous Networks*, 2009, pp. 33–40.

[12] M. Haenggi and R. K. Ganti, "Interference in large wireless networks," *Found. Trends Netw.*, vol. 3, no. 2, pp. 127–248, 2009.

[13] A. Busson, G. Chelius, and J.-M. Gorce, "Interference modeling in csma multi-hop wireless networks," Research Report, INRIA, 6624, 2009.

[14] M. Haenggi, "Mean interference in hard-core wireless networks," *IEEE Communications Letters*, vol. 15, no. 8, pp. 792–794, 2011.

[15] M. Penrose, *Random Geometric Graphs*, 1st ed. New York: Oxford University Press, 2003.

[16] C. Bettstetter, "On the connectivity of ad noc networks," *The computer journal*, vol. 47, pp. 169–181, 2004.

[17] ——, "On the minimum node degree and connectivity of a wireless multihop network," in *ACM international symposium on Mobile ad hoc networking & computing*, 2002, pp. 80–91.

[18] A. Tang, C. Florens, and S. H. Low, "An empirical study on the connectivity of ad hoc networks," in *IEEE Aerospace Conference*, vol. 3, 2003, pp. 1333–1338.

[19] R. Hekmat and P. Van Mieghem, "Connectivity in wireless ad-hoc networks with a log-normal radio model," *Mobile Networks & Applications*, vol. 11, no. 3, pp. 351–360, 2006.

[20] D. Miorandi and E. Altman, "Coverage and connectivity of ad hoc networks presence of channel randomness," in *Proc. IEEE INFOCOM*, vol. 1, 2005, pp. 491–502 vol. 1.

[21] D. Miorandi, "The impact of channel randomness on coverage and connectivity of ad hoc and sensor networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 1062–1072, 2008.

[22] C. Avin, Z. Lotker, F. Pasquale, and Y. A. Pignolet, "A note on uniform power connectivity in the sinr model," *Algorithmic Aspects of Wireless Sensor Networks*, vol. 5804, pp. 116–127, 2009.

[23] E. Lebhar and Z. Lotker, "Unit disk graph and physical interference model: Putting pieces together," in *IEEE International Symposium on Parallel&Distributed Processing*, 2009, pp. 1–8.

[24] G. Mao and B. D. Anderson, "Connectivity of large scale networks: Emergence of unique unbounded component," *submitted to IEEE Transactions on Mobile Computing, available at http://arxiv.org/abs/1103.1991*, 2011.

[25] ——, "Connectivity of large scale networks: Distribution of isolated nodes," *submitted to IEEE Transactions on Mobile Computing, available at http://arxiv.org/abs/1103.1994*, 2011.

[26] A. Hasan and J. G. Andrews, "The guard zone in wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 3, pp. 897–906, 2007.

[27] H. Q. Nguyen, F. Baccelli, and D. Kofman, "A stochastic geometry analysis of dense ieee 802.11 networks," in *Proc. IEEE INFOCOM*, 2007, pp. 1199–1207.

[28] C.-W. Yi, P.-J. Wan, X.-Y. Li, and O. Frieder, "Asymptotic distribution of the number of isolated nodes in wireless ad hoc networks with bernoulli nodes," *IEEE Transactions on Communications*, vol. 54, no. 3, pp. 510–517, 2006.

[29] G. Mao and B. D. O. Anderson, "On the asymptotic connectivity of random networks under the random connection model," in *Proc. IEEE INFOCOM*, 2011, pp. 631–639.

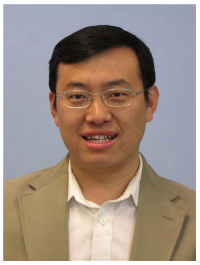
[30] G. Mao and B. D. Anderson, "Towards a better understanding of large scale network models," *accepted to appear in IEEE/ACM Transactions on Networking*, 2011.

[31] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed. New York: Springer, 1999.

[32] T. M. Apostol and M. A. Mnatsakanian, "Sums of squares of distances in m-space," *American Mathematical Monthly*, vol. 110, no. 6, pp. 516–526, 2003.

Tao Yang received her BEng degree and MSc degree in Electrical Engineering from Southwest Jiaotong University, China. She is currently working toward the PhD degree in Engineering at University of Sydney. Her research interests include wireless multihop networks and network performance analysis.

Guoqiang Mao received PhD in telecommunications engineering in 2002 from Edith Cowan University. He joined the School of Electrical and Information Engineering, the University of Sydney in December 2002 where he is a Senior Lecturer now. His research interests include wireless localization techniques, wireless multihop networks, graph theory and its application in networking, and network performance analysis. He is a Senior Member of IEEE and an Associate Editor of IEEE Transactions on Vehicular Technology.



Wei Zhang (S'01–M'06–SM'11) received the Ph.D. degree in electronic engineering from The Chinese University of Hong Kong in 2005. He was a Research Fellow with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, during 2006–2007. From 2008, he has been with the School of Electrical Engineering and Telecommunications, The University of New South Wales, Sydney, Australia, where he is a Senior Lecturer. His current research interests include cognitive radio, cooperative communications, space-time coding, and multiuser MIMO.

He received the best paper award at the 50th IEEE Global Communications Conference (GLOBECOM), Washington DC, in 2007 and the IEEE Communications Society Asia-Pacific Outstanding Young Researcher Award in 2009. He is Co-Chair of Communications Theory Symposium of International Conference on Communications (ICC 2011), Kyoto, Japan. He is an Editor of the IEEE Transactions on Wireless Communications and an Editor of the IEEE Journal on Selected Areas in Communications (Cognitive Radio Series).

Capacity of Interference-limited Three Dimensional CSMA Networks

Tao Yang, Guoqiang Mao
School of Electrical and Information Engineering
The University of Sydney
Email: {tao.yang, guoqiang.mao}@sydney.edu.au

Wei Zhang
School of Electrical Engineering and Telecommunications
The University of New South Wales
Email: wzhang@ee.unsw.edu.au

***Abstract*—In this paper, we study the throughput of interference-limited three dimensional (3D) CSMA networks. Specifically, we consider a network with a total of n nodes uniformly i.i.d. in a cube of edge length $n^{\frac{1}{3}}$. Further, CSMA random access scheme is employed and the SINR model is used to simulate a successful transmission. We first give a sufficient condition on the transmit power required for the CSMA network to be asymptotically almost surely (a.a.s.) connected as $n \to \infty$ under the SINR model. Then, we demonstrate constructively that a throughput of $\Theta\left(\frac{1}{(n \log^2 n)^{\frac{1}{3}}}\right)$ is obtainable by each node for an arbitrarily chosen destination.**

I. Introduction

Wireless multi-hop networks have been increasingly used in military and civilian applications. In many a applications, the region in which the network is deployed is better modeled by a 3D space, instead of a two dimensional (2D) planar area. Examples include a wireless network deployed across different floors inside a building connecting a variety of devices such as computers, smart phones, sensors etc, a network formed by Unmanned Aerial Vehicles and ground devices for reconnaissance and surveillance, and underwater acoustic sensor networks. Capacity of such networks is an important problem. The scaling behavior of capacity when the network becomes sufficiently large is of particular interest.

Existing work on the capacity of wireless multi-hop networks has mainly focused on the analysis of 2D networks [1], [2], including [1] which considered 2D CSMA networks. Limited work has considered the properties of 3D networks where centralized/deterministic scheduling schemes like TDMA are employed [3], [4]. On the other hand, CSMA schemes, which make use of *distributed/randomized* medium access protocols, has become prevailing with widespread adoption. With CSMA, each node checks the status of the wireless channel before sending a packet. If the channel is idle (i.e. no carrier is detected within its carrier-sensing range), then the node starts its transmission, otherwise, defers it, usually by a random amount of time, until the channel becomes idle again. Potential transmitters in the vicinity of an active transmitter are kept off. Wireless signals transmitted at the same time mutually interfere with each other. The SINR (signal to interference plus noise ratio) model has been widely used to capture the impact of interference on the quality of a link and a transmission is considered to be successful iff a minimum SINR requirement has been met [2]. Therefore, it is natural to expect CSMA could improve the network performance by alleviating the interference.

In this paper, we consider a CSMA network with n nodes uniformly i.i.d. in a cube of edge length $n^{\frac{1}{3}}$ and investigate the throughput of the network. The contributions of this paper are:

1) We derive an upper bound on the interference experienced by any receiver in the 3D CSMA network. Using the result, we show that for an arbitrary SINR requirement, there exists a transmission range R_0 such that any two nodes are directly connected if their Euclidean distance is less than or equal to R_0. Based on that, we give a sufficient condition on the transmit power for the CSMA network to be a.a.s. connected under the SINR model as $n \to \infty$. A connected network is a prerequisite for the network to achieve a non-zero throughput.
2) We further show that in the 3D CSMA network, a throughput of $\Theta\left(\frac{1}{(n \log^2 n)^{\frac{1}{3}}}\right)$ is achievable. Compared with the results in [3] and [4], which showed that a throughput of $\Theta\left(\frac{1}{(n \log^2 n)^{\frac{1}{3}}}\right)$ is attainable by using either a deterministic scheduling [3], [4] or without considering the SINR requirement for successful transmissions [4], our result shows that a throughput of $\Theta\left(\frac{1}{(n \log^2 n)^{\frac{1}{3}}}\right)$ is also attainable even when CSMA is used and a minimum SINR is required.

The remainder of this paper is organized as follows: Section II reviews related work; Section III defines the network and metrics being investigated; in Section IV, we give a sufficient condition on the transmit power to have an a.a.s. connected 3D CSMA network; In Section V, we obtain a lower bound on the achievable throughput of 3D CSMA networks; Section VI concludes the paper and discusses future work.

II. Related work

Existing work on studying capacity problem focused mainly on 2D networks. The seminal work [2] showed that in a network with a total of n nodes distributed on a disk of

This work is partially funded by ARC Discovery project: DP110100538. This material is based on research partially sponsored by the Air Force Research Laboratory, under agreement number FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

unit area under the SINR model, the throughput obtainable by each node is 1) $\Theta\left(\frac{1}{\sqrt{n\log n}}\right)$ if nodes are randomly i.i.d. and destination is randomly chosen for each node; 2) $\Theta\left(\frac{1}{\sqrt{n}}\right)$ if nodes' location, traffic pattern and transmission range are optimally arranged. Since this pioneering work, extensive efforts have been made to investigate the capacity in different scenarios. Significant outcomes have been achieved for both static networks [5]–[7] and mobile networks [8], [9]. The paper [9] showed that mobility of nodes can be exploited to significantly improve network capacity at the expense of delay. Other work in the area includes [10], [11] studied the capacity of networks with infrastructure support, and [11] showed that randomly placed base stations can also boost the throughput.

For networks using distributed/random CSMA scheme, the recent work [1] showed that a throughput of $\Theta\left(\frac{1}{\sqrt{n}}\right)$ can be achieved in 2D CSMA networks with randomly chosen destinations. The result is in the same order as the TDMA network considered in [2]. [12], [13] studied the interactions between the transmit power, the carrier-sensing range and the capacity in 2D CSMA networks.

Very limited work (see [3], [4] and references therein) has studied the capacity of 3D networks and all focused on networks employing deterministic schedulings. The paper [3] considered a network deployed in a sphere and showed that a throughput of $\Theta\left(\frac{1}{(n\log^2 n)^{\frac{1}{3}}}\right)$ is feasible. A more recent work [4] studied the capacity of 3D networks under two scenarios, i.e. nodes are regularly placed and nodes are Poissonly distributed.

III. Network models and preliminaries

We consider a network with n nodes uniformly i.i.d. in a cube with edge length $n^{\frac{1}{3}}$.

A. Interference model

Assume all nodes use a common transmit power P. Let $\boldsymbol{x}_k$, $k\in\Gamma$, be the location of node k, where Γ represents the set of indices of all nodes in the network. A transmission from node i to node j is successful iff the SINR at node j is above a threshold β, i.e.

$$\text{SINR}\left(\boldsymbol{x}_i,\boldsymbol{x}_j\right)=\frac{P\ell\left(\boldsymbol{x}_i,\boldsymbol{x}_j\right)}{N_0+\sum\limits_{k\in\mathcal{T}_i}P\ell\left(\boldsymbol{x}_k,\boldsymbol{x}_j\right)}\geq\beta \tag{1}$$

where $\mathcal{T}_i\subseteq\Gamma$ denotes the subset of nodes transmitting at the same time as node i. $\ell\left(\boldsymbol{x}_i,\boldsymbol{x}_j\right)$ represents power attenuation from $\boldsymbol{x}_i$ to $\boldsymbol{x}_j$ and assumes a power-law form, i.e.,

$$\ell\left(\boldsymbol{x}_i,\boldsymbol{x}_j\right)=\left\|\boldsymbol{x}_i-\boldsymbol{x}_j\right\|^{-\alpha} \tag{2}$$

where $\|\cdot\|$ is the Euclidean norm and α is the path-loss exponent. We assume that the background noise N_0 is negligibly small, i.e. $N_0=0$. This assumption is justified because interference is a major factor that weakens performance in wireless networks, while the background noise is typically small and can be combated by increasing the transmit power. As commonly done in the capacity analysis [2]–[4], [9], [11], the impact of small-scale fading is ignored.

Since CSMA typically require an ACK packet to acknowledge a successful transmission, we explicitly consider *bidirectional link* only in the network. In other words, a transmission from node i to node j is successful iff both $\text{SINR}\left(\boldsymbol{x}_i,\boldsymbol{x}_j\right)$ and $\text{SINR}\left(\boldsymbol{x}_j,\boldsymbol{x}_i\right)$ are above β. In that case, we also say that node i and j are directly connected.

B. Definition of throughput

The *channel rate* of a transmission from node i to node j is related to the associated SINR by Shannon theorem, i.e.,

$$R\left(\boldsymbol{x}_i,\boldsymbol{x}_j\right)=B\log_2\left(1+\text{SINR}\left(\boldsymbol{x}_i,\boldsymbol{x}_j\right)\right) \tag{3}$$

where B is the bandwidth of the channel in Hertz. Due to the minimum SINR requirement in (1), the channel rate between a pair of directly connected nodes is at least $B\log_2\left(1+\beta\right)$.

Every node sends data at a rate (bits/sec) to a randomly chosen destination. A node is both a source and a destination node for another node. Therefore the total number of source-destination pairs is n. The per-node *throughput*, denoted by $\lambda\left(n\right)$, is defined as the *maximum* rate that could be achieved by any source-destination pair simultaneously. A throughput of $\lambda\left(n\right)$ is *feasible* if there is a temporal and spatial scheduling scheme such that every node can send $\lambda\left(n\right)$ bits/sec *on average* to its destination, i.e. there exists a sufficiently large positive number τ such that in every finite time interval $\left[\left(j-1\right)\tau,j\tau\right]$ every node can send $\tau\lambda\left(n\right)$ bits to its destination. A throughput is of order $\Theta\left(f\left(n\right)\right)$ bits/sec if there are deterministic constants $0<c<c'<+\infty$ such that $\lim_{n\to\infty}\Pr\left(\lambda\left(n\right)=cf\left(n\right)\text{ is feasible}\right)=1$ and $\lim_{n\to\infty}\Pr\left(\lambda\left(n\right)=c'f\left(n\right)\text{ is feasible}\right)<1$.

C. CSMA random access scheme

In CSMA networks, two nodes, say i and j, are allowed to transmit simultaneously if they can not detect each other's transmission, i.e. both $P\ell\left(\boldsymbol{x}_i,\,\boldsymbol{x}_j\right)$ and $P\ell\left(\boldsymbol{x}_j,\,\boldsymbol{x}_i\right)$ are under a certain detection threshold P_{th} (this is also termed as the *pairwise carrier-sensing decision model* in [1]). This mechanism imposes a *minimum* separation constraint among the concurrent transmitters, known as the *carrier-sensing range*. It readily follows from (2) that the carrier-sensing range R_c is given by

$$R_c=\left(P/P_{\text{th}}\right)^{1/\alpha} \tag{4}$$

Under the carrier-sensing constraint, multiple nodes contend for an opportunity to transmit and at a particular time instant, there can only be one node transmitting in a geographic region determined by R_c. Therefore, the channel rate given by (3) is shared by several nodes in the vicinity over time. Next, we describe how to obtain the *time-average* channel rate (or equivalently the *long-term* channel rate in [1]) for each node.

Same as that in [1], we consider an idealized CSMA scheme. Assume that each node maintains a *countdown timer*, which is initialized to a non-negative random integer. The timer of a node counts down when the node senses the channel idle, otherwise it is frozen. A node initiates its transmission when its timer reaches zero *and* the channel is idle. After finishing transmission, the node resets its timer to a new

random integer. The average countdown time can be distinct for different nodes, which can be set to control the state transition probabilities in the next paragraph.

The above CSMA scheme can be modeled by a Markov chain with state space $\mathbb{S}$, where a state $\mathcal{S} \in \mathbb{S}$ represents the active transmitter set at a particular time instant. A transition between two distinct states $\mathcal{S}, \mathcal{S}' \in \mathbb{S}$ can possibly occur iff $\mathcal{S}' = \{i\} \cup \mathcal{S}$ for $\forall i \in \Gamma$. Transition $\mathcal{S} \to \{i\} \cup \mathcal{S}$ (where $i \notin S$) represents the event that node i will starts its transmission after its timer counts down to zero. Transition $\{i\} \cup \mathcal{S} \to \mathcal{S}$ represents the event that node i finishes its transmission and hence become silent again. Let υ be the set of state transition probabilities and denote the above Markov chain by $\langle \mathbb{S}, \upsilon \rangle$. Then the time-average channel rate available for each node can be characterized by the stationary distribution of $\langle \mathbb{S}, \upsilon \rangle$ [14] [1, Lemma 8].

IV. Connectivity of 3D CSMA Networks

For any throughput to be feasible, a prerequisite is that there exists a path between each pair of source and destination, i.e. the network is connected. In this section, we first derive an upper bound on the interference experienced by any receiver in the network. We further show that for an arbitrarily chosen β, there exists a transmission range R_0 such that a pair of nodes are directly connected if their Euclidean distance is smaller than or equal to R_0. Based on that, we give a sufficient condition on the transmit power for the 3D CSMA network to be a.a.s. connected as $n \to \infty$ under the SINR model.

The following lemma gives an upper bound on the interference.

Lemma 1. *Consider a CSMA network with nodes arbitrarily distributed in a region in $\Re^3$ where the carrier-sensing range is R_c, given by (4). Denote by r_0 the Euclidean distance between an arbitrary receiver and its intended transmitter and $r_0 < R_c$. When the path loss exponent $\alpha > 3$, the maximum interference is upper bounded by $N(r_0)$, where*

$$N(r_0) = 12P(R_c - r_0)^{-\alpha} + 17P\left(\frac{\sqrt{6}}{3}R_c - r_0\right)^{-\alpha} + \frac{9P\left(\left(\frac{251}{6}\alpha^2 - \frac{769}{6}\alpha + 89\right)R_c^2 - 27\sqrt{6}R_c r_0(\alpha - 1) + 54r_0^2\right)}{2\sqrt{6}(\alpha-1)(\alpha-2)(\alpha-3)R_c^3\left(\frac{\sqrt{6}}{2}R_c - r_0\right)^{\alpha-1}} \quad (5)$$

Proof: See Appendix. ∎

Remark 2. Note that the upper bound in Lemma 1 applies to *arbitrary* node distribution in $\Re^3$. Further, the requirement that $\alpha > 3$ is for the interference to be bounded by a constant independent of n. If $\alpha \leq 3$, then the interference given by (13) approaches infinity as $n \to \infty$. In that case, an upper bound on interference can still be found by using the technique presented in the proof of Lemma 1 but that bound will be a function of n. In this paper, we focus on the situation that $\alpha > 3$ to avoid some verbose but straightforward discussion on special cases that occur when $\alpha \leq 3$.

Corollary 3 is a consequence of Lemma 1.

Corollary 3. *Under the same setting as that in Lemma 1, there exists a transmission range $R_0 < R_c$ such that a pair of nodes are directly connected if their Euclidean distance is smaller than or equal to R_0, which is given implicitly by*

$$PR_0^{-\alpha}/N(R_0) = \beta \quad (6)$$

Proof: Noting that $\frac{Pr_0^{-\alpha}}{N(r_0)} \to \infty$ as $r_0 \to 0$, $\frac{Pr_0^{-\alpha}}{N(r_0)} \to 0$ as $r_0 \to R_c^-$ and that $\frac{Pr_0^{-\alpha}}{N(r_0)}$ is a monotonically decreasing function of r_0, therefore there is a unique solution to (6). The rest of the proof is trivial and hence omitted. ∎

Since $P = P_{\text{th}}R_c^{\alpha}$, R_0 in (6) can also be expressed as a function of R_c. Letting $\frac{R_c}{R_0} = x$, (6) can be rewritten as

$$\frac{1}{\beta} = 12(x-1)^{-\alpha} + 17\left(\frac{\sqrt{6}}{3}x - 1\right)^{-\alpha} + \frac{9\left(\left(\frac{251}{6}\alpha^2 - \frac{769}{6}\alpha + 89\right)x^2 - 27\sqrt{6}(\alpha-1)x + 54\right)}{2\sqrt{6}(\alpha-1)(\alpha-2)(\alpha-3)x^3\left(\frac{\sqrt{6}}{2}x - 1\right)^{\alpha-1}} \quad (7)$$

It follows that $R_0 = \frac{R_c}{b}$ and b is the solution to (7), which depends on β and α only. Equation (7) gives a more convenient way to study the relation between P and R_0.

Based on Corollary 3 and the result in [15, Theorem 1] on the connectivity of 3D networks under the unit disk connection model, we obtain the following theorem.

Theorem 4. *Consider a CSMA network with a total of n nodes uniformly i.i.d. in a cube with edge length $n^{\frac{1}{3}}$. Under SINR model, the network is a.a.s. connected as $n \to \infty$ if transmit power*

$$P = P_{th}(b')^{\alpha}(\log n + c(n))^{\frac{\alpha}{3}} \quad (8)$$

where $\lim_{n\to\infty} c(n) = +\infty$, $b' = b(3/4\pi)^{\frac{1}{3}}$ and $\infty > b > 1$ is the solution to (7).

Proof: The theorem readily follows from the result in [15, Theorem 1] with proper scaling and Corollary 3. ∎

An implication of Theorem 4 is that when P is set as that in (8), a.a.s. there exists a temporal and spatial scheduling scheme that allow any pair of nodes in the CSMA network to exchange packets.

V. Feasible Throughput

In this section, we first describing a routing algorithm and then show that a throughput of $\Theta\left(\frac{1}{(n\log^2 n)^{\frac{1}{3}}}\right)$ can be achieved using the routing algorithm in 3D CSMA networks.

Partition the cube into non-overlapping cubelets of edge length $s_n = (4\log n)^{\frac{1}{3}}$. Let X_i be the random number of nodes in a cubelet i. Let $\overline{X} = \max_{i\in\top} X_i$ and $\underline{X} = \min_{i\in\top} X_i$ where $\top$ represents the set of indices of all cubelets. We obtain:

Lemma 5. *As $n \to \infty$, $\Pr\left(\overline{X} \leq c_1 \log n\right) = 1$ and $\Pr(\underline{X} \geq c_2 \log n) = 1$ where $c_1 = 4\left(1 + \frac{\sqrt{3}}{2}\right)$ and $c_2 = 4\left(1 - \frac{\sqrt{2}}{2}\right)$.*

Proof: Note that X_i has a binomial distribution with parameters n and $\frac{s_n^3}{n}$ and $\mathrm{E}[X_i] = s_n^3 = 4\log n$. Using the Chernoff bound, we have that for any $\delta \in (0,1)$, $\Pr[X_i \geq (1+\delta)\mathrm{E}[X_i]] \leq \exp\left(-\frac{\delta^2 \mathrm{E}[X_i]}{3}\right)$ holds.

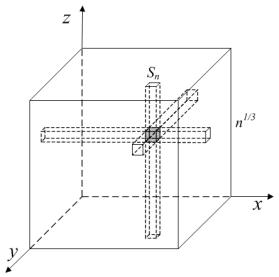


Figure 1. An illustration on the number of routes served by one cubelet.

Let $\delta = \frac{\sqrt{3}}{2}$, then $\Pr\left[X_i \geq 4\left(1+\frac{\sqrt{3}}{2}\right)\log n\right] \leq \frac{1}{n}$. There are a total of $\frac{n}{s_n^3}$ cubelets (here we ignored some trivial discussion on granularity problem caused by $\frac{n}{s_n^3}$ not being an integer). By the union bound, we have $\lim_{n\to\infty}\Pr\left[\bigcup_{i=1}^{n/s_n^3}\left(X_i \geq 4\left(1+\frac{\sqrt{3}}{2}\right)\log n\right)\right] = 0$.

Using a similar method, we have that for any $\delta' \in (0,1)$ $\Pr\left[X \leq (1-\delta')\,\mathrm{E}\left[X\right]\right] \leq \exp\left(-\frac{\delta'^2 \mathrm{E}[X]}{2}\right)$ holds. Taking $\delta' = \frac{\sqrt{2}}{2}$ and using the union bound yields $\underline{X}$. ■

A. The maximum traffic served by each cubelet

For a given β, using (7), (4), we can choose a transmit power so that the transmission range R_0, given by Corollary 3, is $\sqrt{6}s_n$. This value allows any two nodes in two neighboring cubelets to directly communicate with each other under the SINR model. Hence, the channel rate between two nodes, whose Euclidean distance is less than or equal to R_0, is at least $B\log_2(1+\beta)$. Using (7) we can write: $R_c = b\sqrt{6}s_n$, where b is the solution to (7) for the given β.

We employ a similar routing scheme to that used in [4]. For a pair of source and destination nodes located at (x_s, y_s, z_s) and (x_d, y_d, z_d) respectively, packets generated by the source are first relayed to a node closest to (x_s, y_s, z_d), then to a node closest to (x_d, y_s, z_d) and finally delivered to (x_d, y_d, z_d).

As illustrated in Fig. 1, the shaded space represents a cubelet. According to the above routing scheme, only nodes located in the three rectangular cuboid, which are bounded by dashed lines in the figure, possibly need nodes in the shaded cubelet to relay their data. Therefore the maximum number of routes served by each cubelet is

$$N_{\text{routes}} = 3\frac{n^{\frac{1}{3}}}{s_n}\overline{X} = \frac{3\overline{X}}{4^{\frac{1}{3}}}\left(\frac{n}{\log n}\right)^{\frac{1}{3}} \tag{9}$$

Suppose each node sends data at a rate $\lambda(n)$ bits/sec to its destination. The maximum amount of traffic each cubelet *needs* to transmit is $N_{\text{routes}}\lambda(n)$ at most.

B. Time-average channel rate for each node

In this subsection, we first construct a deterministic TDMA scheduling $(\mathcal{S}_t)_{t=1}^m$, where $\mathcal{S}_t \in \mathbb{S}$ is the active transmitter set during time slot t, and determine the time-average channel rate $\boldsymbol{C}_i^{\text{det}}\left[(\mathcal{S}_t)_{t=1}^m\right]$ for node $i \in \Gamma$ under this scheme. Then using the result in [1, Lemma 9], we show that $\boldsymbol{C}_i^{\text{det}}\left[(\mathcal{S}_t)_{t=1}^m\right] \leq \boldsymbol{C}_i^{\text{ran}}\left[\langle\mathbb{S}, \upsilon\rangle\right]$ where $\boldsymbol{C}_i^{\text{ran}}\left[\langle\mathbb{S}, \upsilon\rangle\right]$ is the time-average channel rate for node i under the CSMA scheme described in Section III-C (modeled by the Markov chain $\langle\mathbb{S}, \upsilon\rangle$). Finally we establish a lower bound on $\boldsymbol{C}_i^{\text{ran}}\left[\langle\mathbb{S}, \upsilon\rangle\right]$.

As shown in (3), any pair of directly connected nodes can transmit at a rate at least $B\log_2(1+\beta)$. For convenience, in the following discussion, we consider that the rate equals to $B\log_2(1+\beta)$ and normalize it to 1. We divide time into slots of unit length. It follows that the channel rate available for a particular node i under the scheduling scheme $(\mathcal{S}_t)_{t=1}^m$ is equal to the fraction of time that node i gets to transmit, i.e.

$$\boldsymbol{C}_i^{\text{det}}\left[(\mathcal{S}_t)_{t=1}^m\right] = \frac{1}{m}\sum_{t=1}^{m}\mathbf{1}\left(i \in \mathcal{S}_t\right)$$

We group adjacent cubelets into non-overlapping cubes and each cube contains $(k+1)^3$ cubelets, where $k = \left\lceil b\sqrt{6}\right\rceil$ so that $ks_n \geq R_c = b\sqrt{6}s_n$. Using Lemma 5, a.a.s. there are at most $c_1\log n$ nodes in every cubelet. Based on the above discussion, a deterministic scheduling algorithm can be designed such that within time slots from $t=1$ to $t=(k+1)^3 c_1\log n$, each node gets at least one time slot to transmit while the set of concurrent transmitters meets the CSMA constraints. Denote by $\mathcal{S}_t$ the concurrent transmitter set during time slot t. It follows that $\mathcal{S}_t \in \mathbb{S}$ for $t \in \left[1, (k+1)^3 c_1\log n\right]$ and the fraction of time spent on each $\mathcal{S}_t$, $t \in \left[1, (k+1)^3 c_1\log n\right]$ is $\frac{1}{(k+1)^3 c_1\log n}$. Letting $m = (k+1)^3 c_1\log n$, it then follows that there is a deterministic scheduling that can achieve a time-average channel rate of at least $\frac{B\log(1+\beta)}{(k+1)^3 c_1\log n}$ for node i, i.e.

$$\boldsymbol{C}_i^{\text{det}}\left[(\mathcal{S}_t)_{t=1}^m\right] \geq \frac{B\log_2(1+\beta)}{(k+1)^3 c_1\log n} \tag{10}$$

Using the result in [1, Lemma 9] which states that there exists a properly designed CSMA scheme that delivers suitable state transition probabilities υ, such that for each node i, the following holds

$$\boldsymbol{C}_i^{\text{ran}}\left[\langle\mathbb{S}, \upsilon\rangle\right] \geq \boldsymbol{C}_i^{\text{det}}\left[(\mathcal{S}_t)_{t=1}^m\right] \tag{11}$$

where $\boldsymbol{C}_i^{\text{det}}\left[(\mathcal{S}_t)_{t=1}^m\right]$ in (11) is the time-average channel rate available for node i under a deterministic scheduling scheme.

Combining (11), (10) and (3), it can be established that for node $i \in \Gamma$, $\boldsymbol{C}_i^{\text{ran}}\left[\langle\mathbb{S}, \upsilon\rangle\right] \geq \frac{B\log_2(1+\beta)}{(k+1)^3 c_1\log n}$.

Lemma 5 also tells that the minimum number of nodes in every cubelet is greater than or equal to $c_2\log n$ a.a.s. Therefore, the minimum time-average channel rate available for each cubelet under CSMA scheme is

$$\frac{(c_2\log n)\,B\log_2(1+\beta)}{(k+1)^3 c_1\log n} = \frac{c_2 B\log_2(1+\beta)}{c_1(k+1)^3} \tag{12}$$

C. Lower bound on throughput

For *any* per-node throughput $\lambda(n)$ to be feasible, the traffic load for each cubelet should not exceed the time-average channel rate available for each cubelet, i.e.,

$$N_{\text{routes}}\lambda(n) \leq \frac{c_2 B\log_2(1+\beta)}{c_1(k+1)^3} \text{ bits/sec}$$

which results in a lower bound on the feasible per-node throughput. This is summarized in the Theorem 6, which forms another major contribution of this paper.

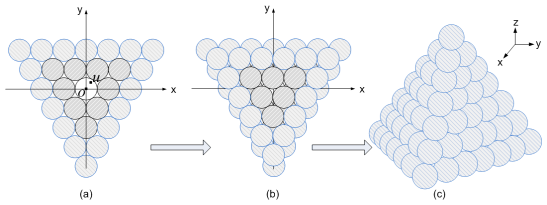


Figure 2. Densest sphere packing.

Theorem 6. *For the considered CSMA networks, there exists a deterministic constant $c > 0$, independent of α, P_{th} and β, such that a per-node throughput $\lambda(n) = \frac{c\log_2(1+\beta)}{(k+1)^3}\left(\frac{1}{(n\log^2 n)^{\frac{1}{3}}}\right)$ bits/sec is feasible a.a.s. as $n \to \infty$, where $k = \lceil b\sqrt{6} \rceil$ and b is the solution to (7) for a given β.*

VI. Conclusion

In this paper, we studied the throughput of 3D CSMA networks. We first provided a sufficient condition on the transmit power for having an a.a.s. connected 3D CSMA network under the SINR model. Then, using a simple routing scheme, we obtained a per-node throughput of $\Theta\left(\frac{1}{(n\log^2 n)^{\frac{1}{3}}}\right)$ is feasible even when distributed/random access scheme is used and a minimum SINR for each successful transmission is specified. It remains our future work to study the optimum access scheme that maximizes the per-node throughput.

Appendix: Proof of Lemma 1

The derivation of the upper bound on interference is similar to that used in [16]. The difference is in that, here we use densest sphere packing in 3D space to derive the upper bound.

Construct a coordinate system such that the origin o is at a transmitter w. Consider a transmission from w to its receiver u located at $\boldsymbol{u}$ and define $r_0 \triangleq \|\boldsymbol{u}\|$.

Draw a sphere of radius $R_c/2$ centered at each concurrent transmitter. The two spheres centered at two closest transmitters cannot overlap. The maximum interference happens when these spheres are placed in the densest way, which is to place the sphere centers at the vertices of a *face-centered cubic lattice* [17, p.9]. See Fig. 2 for an illustration.

Group the sphere centers into tiers of increasing distance from the origin. The picture (a) in Fig.2 shows the spheres in the 1st tier (in dark shade) and the 2nd tier (in light shade) whose centers are located on the $x-y$ plane. All of the spheres in the 2nd tier whose centers are above the $x-y$ plane (including those on the $x-y$ plane) are shown in light shade in picture (c). Although Fig.2 only shows the packing along $+z$ axis, the packing goes along $-z$ axis as well in the same way as along $+z$ axis. The number of interferers in the j^{th} tier is $27j^2+2$. Let $\boldsymbol{x}_i$ be the location of an interferer. The minimum distance from an interferer in the j^{th} tier to the origin is $\frac{\sqrt{6}}{3}jR_c$. Denote by $I(r_0)$ the interference at node $\boldsymbol{z}$. Since $\|\boldsymbol{x}_i - \boldsymbol{u}\| \geq \|\boldsymbol{x}_i\| - r_0$, (2) we obtain

$$I(r_0) \leq \frac{12P}{(R_c - r_0)^{\alpha}} + \frac{17P}{\left(\frac{\sqrt{6}}{3}R_c - r_0\right)^{\alpha}} + \sum_{j=2}^{\infty} \frac{P\left(27j^2+2\right)}{\left(\frac{\sqrt{6}}{3}jR_c - r_0\right)^{\alpha}} \quad (13)$$

The first two items in RHS of (13) account for the interference caused by the 1st tier interferers. Let U_j, $j = 2, \ldots, \infty$, be random variables uniformly and i.i.d. in $\left[j-\frac{1}{2}, j+\frac{1}{2}\right]$. It follows from the convexity of $\left(27j^2+2\right)\left(\frac{\sqrt{6}}{3}jR_c - r_0\right)^{-\alpha}$ when $j \geq 2, \alpha > 3$ and Jensen's inequality that

$$\begin{aligned}
& P\sum_{j=2}^{\infty}\left(27j^2+2\right)\left(\frac{\sqrt{6}}{3}jR_c - r_0\right)^{-\alpha} \\
= \;& P\sum_{j=2}^{\infty}\left(27\mathrm{E}\left(U_j\right)^2+2\right)\left(\frac{\sqrt{6}}{3}\mathrm{E}\left(U_j\right)R_c - r_0\right)^{-\alpha} \\
\leq \;& P\sum_{j=2}^{\infty}\mathrm{E}\left(\left(27U_j^2+2\right)\left(\frac{\sqrt{6}}{3}U_jR_c - r_0\right)^{-\alpha}\right) \\
= \;& P\sum_{j=2}^{\infty}\int_{j-1/2}^{j+1/2}\left(27x^2+2\right)\left(\frac{\sqrt{6}}{3}xR_c - r_0\right)^{-\alpha}dx \\
= \;& \frac{9P\left(\left(\frac{251}{6}\alpha^2 - \frac{769}{6}\alpha + 89\right)R_c^2 - 27\sqrt{6}R_cr_0\left(\alpha-1\right) + 54r_0^2\right)}{2\sqrt{6}\left(\alpha-1\right)\left(\alpha-2\right)\left(\alpha-3\right)R_c^3\left(\frac{\sqrt{6}}{2}R_c - r_0\right)^{\alpha-1}} \quad (14)
\end{aligned}$$

Substituting (14) into (13), Lemma 1 is proved.

References

[1] C.-K. Chau, M. Chen, and S. C. Liew, "Capacity of large-scale csma wireless networks," *IEEE/ACM Trans. Netw.*, vol. 19, no. 3, pp. 893–906, 2011.

[2] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[3] ——, "Internets in the sky: capacity of 3d wireless networks," in *Proc. IEEE CDC*, 2000.

[4] P. Li, M. Pan, and Y. Fang, "The capacity of three-dimensional wireless ad hoc networks," in *Proc. IEEE INFOCOM*, 2011.

[5] O. Dousse, M. Franceschetti, and P. Thiran, "On the throughput scaling of wireless relay networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2756–2761, 2006.

[6] M. Franceschetti, O. Dousse, D. N. C. Tse, and P. Thiran, "Closing the gap in the capacity of wireless networks via percolation theory," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 1009–1018, 2007.

[7] A. Ozgur, O. Leveque, and D. N. C. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3549–3572, 2007.

[8] G. Sharma, R. Mazumdar, and B. Shroff, "Delay and capacity trade-offs in mobile ad hoc networks: A global perspective," *IEEE/ACM Trans. Netw.*, vol. 15, no. 5, pp. 981–992, 2007.

[9] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477–486, 2002.

[10] P. Li and Y. Fang, "Impacts of topology and traffic pattern on capacity of hybrid wireless networks," *IEEE Trans. Inf. Theory*, vol. 8, no. 12, pp. 1585–1595, 2009.

[11] J.-w. Cho, S.-L. Kim, and S. Chong, "Capacity of interference-limited ad hoc networks with infrastructure support," *IEEE Commun. Letters*, vol. 10, no. 1, pp. 16–18, 2006.

[12] T.-S. Kim, H. Lim, and J. C. Hou, "Understanding and improving the spatial reuse in multihop wireless networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 10, pp. 1200–1212, 2008.

[13] T.-Y. Lin and J. C. Hou, "Interplay of spatial reuse and sinr-determined data rates in csma/ca-based, multi-hop, multi-rate wireless networks," in *Proc. IEEE INFOCOM*, 2007.

[14] X. Wang and K. Kar, "Throughput modelling and fairness issues in csma/ca based ad-hoc networks," in *Proc. IEEE INFOCOM*, 2005.

[15] V. Ravelomanana, "Extremal properties of three-dimensional sensor networks with applications," *IEEE Trans. Mobile Comput.*, vol. 3, no. 3, pp. 246–257, 2004.

[16] T. Yang, G. Mao, and W. Zhang, "Connectivity of wireless csma multi-hop networks," in *Proc. IEEE ICC*, 2011.

[17] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed. New York: Springer, 1999.

On the Information Propagation Process in Mobile Vehicular Ad Hoc Networks

Zijie Zhang, *Student Member, IEEE*, Guoqiang Mao, *Senior Member, IEEE*, and Brian D. O. Anderson, *Life Fellow, IEEE*

***Abstract*—In this paper, we study the information propagation process in a 1-D mobile ad hoc network formed by vehicles Poissonly distributed on a highway and traveling in the same direction at randomly distributed speeds that are independent between vehicles. Considering a model in which time is divided into time slots of equal length and each vehicle changes its speed at the beginning of each time slot, independent of its speed in other time slots, we derive analytical formulas for the fundamental properties of the information propagation process and the information propagation speed (IPS). Using the formulas, one can straightforwardly study the impact on the IPS of various parameters such as radio range, vehicular traffic density, and time variation of vehicle speed. The accuracy of the results is validated using simulations. The research provides useful guidelines on the design of vehicular ad hoc networks (VANETs).**

***Index Terms*—Information propagation speed (IPS), mobile ad hoc network, vehicular ad hoc network (VANET).**

I. Introduction

A VEHICULAR ad hoc network (VANET) is a mobile multihop network formed by vehicles traveling on the road. As a new way of communication, VANETs have attracted significant interest in not only academia but in industry as well [1]. IEEE has taken up working on new standards for VANETs, such as the IEEE 1609 Family of Standards for Wireless Access in Vehicular Environments (WAVE) [2]. Furthermore, there are many projects on VANETs such as InternetITS in Japan, Network on Wheels in Germany, and the PReVENT project in Europe [3]. In this paper, we study the expected propagation speed for a piece of information to be broadcast along the road in a VANET, which is referred to as the *information propagation speed (IPS)*. Due to the mobility of vehicles, the topology of a VANET is changing over time. Furthermore, traffic density on a road can significantly vary, depending on the time of day or day of the week. Therefore, the information propagation process in a VANET can be quite different from that in a static network.

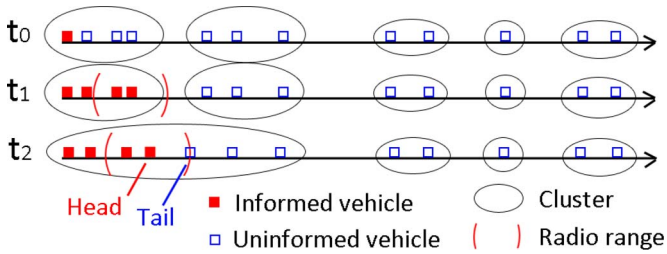


Fig. 1. Topology of a VANET at different time instances. In the figures in this paper, the positive direction of the axis is the direction of information propagation.

A VANET is usually partitioned into a number of clusters [4], [5], where a *cluster* is a maximal set of vehicles in which every pair of vehicles is connected by at least one multihop path. Due to the mobility of vehicles, the clusters are splitting and merging over time. Therefore, the information propagation in a VANET is typically based on a store-and-forward scheme, same as that in a delay-tolerant network [4]. Considering the example shown in Fig. 1, a piece of information starts to propagate from the origin toward the positive direction of the axis at time t_0. The vehicles that have received this piece of information are referred to as the *informed vehicles*, where other vehicles are *uninformed*. As indicated by the leftmost ellipse, the first informed vehicle is inside a cluster of four vehicles at time t_0. At time t_1, the message is forwarded, in a multihop manner, to the foremost vehicle in its cluster. The propagation of the message within a cluster, which begins at t_0 and ends at t_1, is called a *forwarding process*. In a forwarding process, the IPS is determined by the per-hop delay and the length of the cluster. The *per-hop delay* β is the time required for a vehicle to receive and process a message before it is available for further retransmission [6]. The value of β depends on the practical implementation, and a common assumption for the value of β reflecting typical technology is 4 ms [6]. We show that the

Manuscript received July 7, 2010; revised January 10, 2011 and March 25, 2011; accepted April 1, 2011. Date of publication April 21, 2011; date of current version June 20, 2011. This work was supported in part by ARC Disk Overy Project DP110100538 and in part by the Air Force Research Laboratory, under Agreement FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government. The review of this paper was coordinated by Prof. Y. Cheng.

Z. Zhang is with the School of Electrical and Information Engineering, University of Sydney, Sydney, N. S. W. 2006, Australia (e-mail: zijie.zhang@sydney.edu.au).

G. Mao is with the School of Electrical and Information Engineering, University of Sydney, Sydney, N. S. W. 2006, Australia, and also with National Information and Communications Technology (ICT) Australia, Sydney, N. S. W. 2000, Australia (e-mail: guoqiang.mao@sydney.edu.au).

B. D. O. Anderson is with the Research School of Information Sciences and Engineering, Australian National University, Canberra, A. C. T. 0200, Australia, and also with National ICT Australia, Canberra, A. C. T. 0200 Australia (e-mail: brian.anderson@anu.edu.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2011.2145012

per-hop delay has a significant impact on the IPS, particularly when the vehicle density is high.

Define the *head* at time t to be the informed vehicle with the largest coordinate at time t. Define the *tail* at time t to be the uninformed vehicle with the smallest coordinate at time t. Two vehicles can directly communicate with each other if and only if their Euclidean distance is smaller than or equal to the radio range r_0. Although this so-called *unit disk model* is a simplified model, it can be indicative for real-world scenarios. A realistic radio model usually takes into account statistical variations of the received signal power around its mean value [7]. It is shown in [7] that these variations can actually increase the connectivity of a network. Therefore, the analysis under the unit disk model provides a conservative estimate on the performance of a VANET. (We further explore this issue in Section VIII-C.) As shown in Fig. 1, at time t_1, the tail is outside the radio range of the head. Then, a *catch-up process* begins, during which the informed vehicles hold the information until the head catches up with the tail (at time t_2). We study both the forwarding process and the catch-up process in this paper.

The main contributions of this paper are given as follows: 1) Analytical results on the distribution of the time required for a catch-up process are provided. The impact of vehicle density, vehicle speed distribution, and vehicle speed variation over time is considered, where previous research (e.g., [5] and [8]) failed to consider the impact of the time-varying vehicle speed, which results in an unrealistic model according to traffic theory [9]. 2) A first passage phenomenon, which will be introduced later, is considered in the analysis. This first passage phenomenon is a major technical hurdle in the accurate analysis of the catch-up process when vehicle speeds are allowed to change over time. To the best of our knowledge, this work is the first attempt to consider the impact of the first passage phenomenon on the information propagation in VANETs. 3) The forwarding process is studied, taking into account the per-hop propagation delay and packet collision, which give an upper bound on the IPS that was not recognized by previous research (e.g., [5], [8], and [10]). It is shown that the per-hop delay and packet collision have significant impacts on the IPS, particularly when the vehicle density is high. Finally, a closed-form equation is derived for the distribution of the length of a cluster, where, in previous research, only numerical solutions [3] or approximate results [5] were provided. Based on the preceding results, the analytical results for the IPS are derived, including the impact of various parameters such as radio range, vehicular traffic density, and the time variation of vehicle speed. The results in this paper provide useful guidelines on the design of a mobile VANET.

The rest of this paper is organized as follows: Section II reviews the related work. Section III introduces the mobility model and network model used in this paper. The analysis on the catch-up process for a generic speed distribution is given in Section IV, followed by the analysis on the catch-up process for the Gaussian speed distribution in Section V. In Section VI, the analysis on the forwarding process is provided, including the analysis on the cluster length. Based on the preceding results, the IPS is derived in Section VII. Finally, Section IX concludes this paper.

II. Related Work

In recent years, VANETs have attracted significant interest due to their large number of potential applications [1]. In [11], Fracchia and Meo introduced the design of a warning delivery service in VANETs. They studied the propagation of a warning message in a 1-D VANET, where vehicles move in the opposite direction from the propagating direction of the warning message. However, their analysis is based on an oversimplified assumption that the topology does not change over time during the information propagation process. In [12], Camara *et al.* studied the propagation speed of public safety warning messages in a VANET when infrastructures (road side units) are destroyed by natural disasters such as flooding and earthquakes. Through simulation, they showed that using vehicles as virtual road side units can significantly speed up the warning message distribution process, compared with traditional emergency alert systems that rely on infrastructures.

The IPS is an important performance metric for VANETs, particularly for the safety messaging applications [12], [13]. In [6], Wu *et al.* studied the IPS through simulations. They used a commercial microscopic traffic simulator, i.e., CORSIM [14], to simulate the traffic on a highway. Then, the topology data from CORSIM were imported into a wireless communication simulator to study the properties of the information propagation process. They showed that the IPS significantly varies with vehicle densities.

There are analytical studies on VANETs based on the assumption that the vehicle speed does not change over time, which is referred to as the *constant speed model*. In [3], Yousefi *et al.* provided analytical results on the distribution of intervehicle distance in a 1-D VANET under the constant-speed model and the Poisson arrival model: In the *Poisson arrival model*, the number of vehicles passing an observation point on the road during any time interval follows a homogeneous Poisson process with intensity λ. By applying the results used in studying the busy period in queuing theory, they further analyzed the connectivity distance, which is a quantity similar to the cluster length to be introduced later in this paper. However, they did not provide a closed-form formula for the distribution of the cluster length. In [8], Agarwal *et al.* studied the IPS in a 1-D VANET where vehicles are Poissonly distributed and move at the same speed but in either the positive or the negative direction of the axis. They derived the upper and lower bounds for the IPS, which provided a hint on the impact of vehicle density on the IPS. However, the bounds are not tight, and many factors, e.g., time variation of speeds and propagation delay, were ignored in their analysis. In [5], Wu *et al.* considered a 1-D VANET where vehicles are Poissonly distributed and the vehicle speeds are uniformly distributed in a designated range. They provided a numerical method to compute the IPS under two special network models, i.e., when the vehicle density is either very low or very high, which are obviously oversimplified models [4], [9]. The aforementioned studies [3], [5], [8] were all based on the constant speed model. In this paper, the time variation of vehicle speed is considered and is shown to have a significant impact on the information propagation process in VANETs.

III. System Model

A. Mobility Model

A synchronized random walk mobility model is considered in this paper. Specifically, time is divided into time slots of equal length τ. Each vehicle randomly chooses its new speed at the beginning of each time slot, independent of other vehicles and its own speed in other time slots, according to a certain distribution with a mean value $E[v]$. It is shown later that the constant-speed model forms a special case in the aforementioned mobility model when $\tau \to \infty$. Due to a limited speed acceleration, the vehicle speed in the real world does not change as rapidly as in the aforementioned mobility model. Therefore, as will be shown in Section VIII-C, the results based on the aforementioned model provide a conservative estimate on the IPS, which is desirable for the safety messaging applications (e.g., [11] and [12]). Furthermore, we also discuss the impact of a nonsynchronized mobility model at the beginning of Section V.

In the preceding model, the speed of a vehicle can be considered as having a constant component $E[v]$ and a variable component with zero mean. Accordingly, the vehicular network can be decomposed into two components: 1) a network in which all vehicles travel at a constant speed and 2) a network in which vehicles travel at speeds following the same prescribed distribution $f_v(v)$ with zero mean. Our analysis focuses on the IPS in the second network component, where $f_v(v)$ is the probability density function (pdf) of the speed distribution. The first network component is separately considered and is combined into the result at the end of the analysis. Furthermore, a positive (negative) value of the speed means that the vehicle is traveling in the same (opposite) direction as the direction of information propagation. Therefore, when $E[v]$ is a positive (negative) value, our analysis provides the results for a VANET in which a message propagates in the same (opposite) direction as the direction of the vehicle traffic flow. In this paper, the analysis is first performed for a generic speed distribution. Then, detailed analytical results are given for the Gaussian speed distribution with standard deviation σ, which is commonly used for the VANETs on a highway [3], [9], [15].

The aforementioned speed-change time interval τ depends on practical conditions, e.g., a sports car may more frequently change its speed than a heavy truck. Reasonable values for the time interval can be from 1 to 25 s [16]. The vehicle mobility parameters, i.e., $E[v]$, σ, and τ, are taken from practical measurements. Typical values for $E[v]$ and σ are given in [15], where the usual record time intervals for a vehicle speed monitor are $\tau = 1$ s, 5 s [17]. We conduct our analysis in the discrete-time domain ($t = i\tau$) to obtain closed-form analytical equations, which give better insight into the impact of different parameters on the IPS. Extension to the continuous-time domain is straightforward, following the procedure outlined in this paper.

B. Network Model

We adopt a commonly used traffic model in vehicular traffic theory [9] in which vehicles independently travel in the same direction on a 1-D infinite line and follow the Poisson arrival model with a rate λ veh/s. The Poisson arrival model is a commonly used traffic model in vehicular traffic theory based on real-world measurements [9]. Furthermore, the Poisson arrival model and the Poisson distribution of the vehicles are also commonly used traffic models in studies of VANETs [3], [8], [11].

The following lemma relates the spatial distribution of the vehicles on the road to the Poisson arrival model of the vehicles. The result on the spatial distribution of vehicles is used in the rest of this paper.

Lemma 1: If the traffic in a VANET follows the Poisson arrival model with rate λ and the speed of each vehicle changes at the beginning of each time slot, independent of other vehicles, according to $f_v(v)$, then at any time instant, the spatial distribution of the vehicles on the road follows a homogeneous Poisson process with intensity $\rho = \lambda \int_{-\infty}^{\infty}(f_v(v)/v)dv$.

Proof: It has been shown in [3] that, if the vehicle speeds do not change over time, then at any time instant, the distances between adjacent vehicles (intervehicle distance l) are independent and follow an exponential distribution with rate parameter $\rho = \lambda \int_{-\infty}^{\infty}(f_v(v)/v)dv$. It follows that the spatial distribution of the vehicles follows a homogeneous Poisson process with intensity ρ.

Next, we apply the mathematical induction to study the spatial distribution of the vehicles under our mobility model, in which vehicles are allowed to change their speeds from one time slot to another. In the first time slot [0, τ), it is straightforward to show that the spatial distribution of the vehicles follows a homogeneous Poisson process with intensity ρ, because the speed does not change during a time slot.

Assume that, in the ith time slot $[(i-1)\tau, i\tau)$, the spatial distribution of the vehicles still follows a homogeneous Poisson process with intensity ρ. When the next time slot begins, each vehicle chooses its new speed, independent of other vehicles, according to $f_v(v)$. Then, according to the random splitting property of a Poisson process [18], the spatial distribution of the vehicles traveling at the speed $v \in [v_m, v_m + dv_m)$ during the $i+1$th time slot $[i\tau, (i+1)\tau)$ follows a homogeneous Poisson subprocess with intensity $\rho f_v(v_m)dv_m$. These vehicles have the same speed so that, at any time instant during the $i+1$th time slot, the spatial distribution of these vehicles does not change. Then, according to the random coupling property of a Poisson process [18], the spatial distribution of all the vehicles in the $i+1$th time slot follows a homogeneous Poisson process, which is the sum of all the subprocesses, with the intensity $\int_{-\infty}^{\infty} \rho f_v(v_m)dv_m = \rho$. ■

Note that previous research [3] only considered that the vehicle speeds do not change over time. In Lemma 1, we consider the time variation of vehicle speeds in the analysis of vehicle distribution on the road.

IV. Catch-Up Process for a Generic Speed Distribution

In this section, we study the catch-up process in a VANET where the vehicle speeds follow a generic pdf $f_v(v)$. Without loss of generality, it is assumed that the catch-up process starts at time 0. The *displacement* x of a vehicle at time t is defined

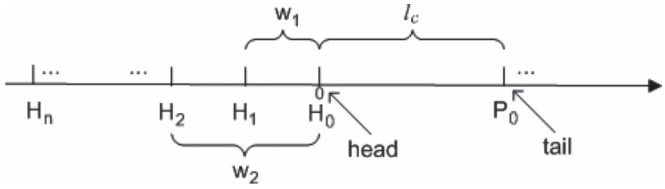


Fig. 2. VANET at the beginning of a catch-up process with gap l_c, where l_c is the Euclidean distance between the head and the tail at time 0. Hereinafter, a catch-up process where the distance between the head and the tail is l_c at time 0 is referred to as a *catch-up process with gap* l_c.

as the difference between the position of the vehicle at time 0 and its position at time t.

A. Modeling the Movement of a Single Vehicle

Denote by $p(x,t)$ the probability that the displacement of a vehicle is x at time t. Because the speed does not change during a time slot, $p(x,\tau)$ can be easily obtained from $f_v(v)$.

Due to the independence of the vehicle speeds in different time slots (hence, the displacements), we have, for $t = i\tau$

$$p(x,t) = p(x,i\tau) = \overbrace{(p * p * \cdots * p)(x,\tau)}^{\text{i–fold convolution}}. \quad (1)$$

The calculation of the aforementioned i-fold convolution can be simplified by using the Fourier and inverse Fourier transform.

B. Modeling the Movement of the Head and the Tail

Denote by H_m (P_m) the mth vehicle to the left of the head H_0 (to the right of the tail P_0) at time 0, as shown in Fig. 2. Define w_m to be the Euclidean distance between H_m and H_0 at time 0.

Let us first consider the movement of the head. Define $x_m(t)$ to be the displacement of H_m at time t. Define $y(t)$ to be the displacement of the head at time t. Note that the head vehicle at time 0 is not necessarily the head vehicle at time t because the original head may be overtaken by another informed vehicle during time $(0,t]$. It follows that

$$y(t) = \max\{x_0(t), x_1(t) - w_1, x_2(t) - w_2, \ldots, x_n(t) - w_n\} \quad (2)$$

where n is the number of vehicles to the left of the head that have the potential to overtake the head vehicle at time 0.

Because the movement of a vehicle is independent of other vehicles, $x_m(t)$ and $x_j(t)$ are independent for any $m \neq j$. Therefore, the cumulative distribution function (cdf) of the displacement of the head at time t is

$$\begin{aligned}\Pr(y(t) \le y) &= \prod_{m=0}^{n} \Pr(x_m(t) - w_m \le y) \\ &= \prod_{m=0}^{n} \left(\int_0^{\infty} \int_{-\infty}^{y+w_m} p(x,t) f_{w_m}(w_m)\, dx\, dw_m \right) \quad (3)\end{aligned}$$

where $p(x,t)$ is given by (1), w_m is the distance between H_m and H_0 at time 0, $w_0 = 0$, and $f_{w_m}(w_m)$ is the pdf of w_m.

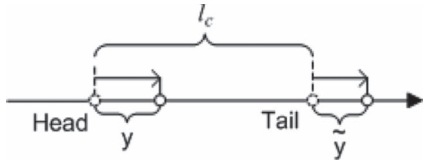


Fig. 3. Displacements of the head and the tail at time t in a catch-up process with gap l_c. The reduction of distance is $z = y - \tilde{y}$.

As an easy consequence of the Poisson distribution of the vehicles (proved in Lemma 1), the intervehicle distance follows an exponential distribution. Therefore, we have

$$f_{w_m}(w_m) = \frac{\rho e^{-\rho w_m} (\rho w_m)^{m-1}}{(m-1)!}, \text{ for } m \ge 1. \quad (4)$$

Define $p_h(y,t)$ to be the probability that the displacement of the head is y at time t. Then

$$p_h(y,t) = \frac{\partial \Pr(y(t) \le y)}{\partial y}. \quad (5)$$

The calculation is, however, tedious for a generic speed distribution. Therefore, only the methodology for the analysis of the catch-up process is shown in this section. A detailed analytical result for the catch-up process under the Gaussian speed distribution is shown in Section V.

Denote by $p_g(\tilde{y},t)$ the probability that the displacement of the tail is $\tilde{y}$ at time t. The analysis for the movement of the tail is similar to that for the head and is, therefore, omitted.

C. Catch-Up Delay

As shown in Fig. 2, consider a catch-up process where the Euclidean distance between the head and the tail is l_c at the beginning of the catch-up process (time 0), which is referred to as *a catch-up process with gap* l_c. Define the *catch-up delay* t_c to be the time taken from time 0 until the time when the head and the tail move into the radio range of each other for the first time, i.e., $t_2 - t_1$ in Fig. 1. We do not consider the rare event that the distance between the head and the tail becomes larger than r_0 again during $(t_c, t_c + \beta)$, which may cause the transmission of a packet being interrupted, because the per-hop delay β (e.g., 4 ms [6]) is usually much smaller than the time interval for a vehicle to change speed (typically longer than a second [16]). It is worth noting that there is a first passage phenomenon in the catch-up process, i.e., the catch-up process finishes as soon as the head and the tail move into the radio range of each other. Therefore, the catch-up delay is t_c *if* and *only if* (iff) the distance between the head and the tail reduces from l_c at time 0 to r_0 *for the first time* at time t_c. This first passage phenomenon is essential for the analysis of the catch-up process.

Denote by $p_H(z,t)$ the probability that the reduction of the Euclidean distance between the head and the tail is z at time t, with regard to their original distance at time 0. As shown in Fig. 3, it can be shown that

$$p_H(z,t) = \int_{-\infty}^{\infty} p_h(y,t) p_g(y - z, t) dy. \quad (6)$$

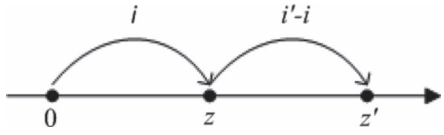


Fig. 4. Cass of random walks taking i' time slots to walk from 0 to z' through an intermediate point z.

Note that the aforementioned equation can be converted to convolution if $p_g(\tilde{y},t)=p_g(-\tilde{y},t)$, which is the case to be introduced in the next section when the vehicle speed follows a Gaussian distribution.

Denote by $h(z,i)$ the probability that the reduction in the distance between the head and the tail reaches z in the ith time slot $[(i-1)\tau, i\tau)$. Therefore

$$h(z,i)=\int_{(i-1)\tau}^{i\tau} p_H(z,t)dt. \tag{7}$$

To obtain a closed-form result, the probability of the reduction in the distance between the head and the tail being z at time $t\in[(i-1)\tau,i\tau)$ is considered to be approximately equal to the probability of the reduction of the distance between the head and the tail being z at time $t=i\tau$. This approximation provides a fairly accurate result when τ is small (e.g., $\tau=1$ s, 5 s), as shown in Sections VIII-A and C. Therefore, from (7)

$$h(z,i)=\int_{(i-1)\tau}^{i\tau} p_H(z,t)dt\approx \tau p_H(z,i\tau). \tag{8}$$

Define $\xi(z,i)$ to be the first passage probability [19] of $h(z,i)$, viz., the probability that the reduction in the distance between the head and the tail reaches z in the ith time slot $[(i-1)\tau,i\tau)$ *for the first time* since time 0. The relationship between $\xi(z,i)$ and $h(z,i)$ can be studied as the first passage time in a stochastic process [20]. The first passage time of a diffusing particle or a random walker is the time at which the particle or the random walker *first* reaches a specified site.

A standard procedure is applied to determine the first passage probability $\xi(z,i)$ [19], [20]. As shown in Fig. 4, consider a class of random walks starting at time 0, and walking from point 0 to z' must proceed by going through a point z. The transition from 0 to z' can be decomposed into two independent stages: in the first stage, an agent walks from 0 to z for the first time in the ith time slot; in the second stage, the agent walks from z to z' in the $i'-i$th time slot, not necessarily for the first time. Then, we can obtain the following [19], [20]:

$$h(z',i')=\sum_{i=1}^{i'}\xi(z,i)h(z'-z,i'-i). \tag{9}$$

The convolution can be simplified by the Z-transform with regard to i, which is denoted by the operator $\mathcal{Z}$. According to the convolution theorem, (9) becomes

$$(\mathcal{Z}h)(z',s)=(\mathcal{Z}\xi)(z,s)(\mathcal{Z}h)(z'-z,s) \tag{10}$$

$$\text{and thus}\quad (\mathcal{Z}\xi)(z,s)=\frac{(\mathcal{Z}h)(z',s)}{(\mathcal{Z}h)(z'-z,s)}. \tag{11}$$

Then, by inverse Z-transform, we can obtain $\xi(z,i)$. Denote by $F_\xi(z,i)$ the cdf of $\xi(z,i)$ with regard to i, i.e., the probability that the reduction in the distance between the head and the tail has reached z during time $(0,i\tau]$. It follows that $F_\xi(l_c-r_0,i)$ is the probability that the head and the tail have moved into the radio range of each other during time $(0,i\tau]$. Therefore, the expected catch-up delay (t_c) for a catch-up process with gap l_c is

$$E[t_c|l_c]=\tau\sum_{i=1}^{\infty}\left(1-F_\xi(l_c-r_0,i)\right). \tag{12}$$

D. Distribution of the Gaps l_c

Denote by $f_l(l)$ the pdf of the Euclidean distance between any two adjacent vehicles. Due to the Poisson distribution of vehicles, it is evident that $f_l(l)=\rho e^{-\rho l}$. Denote by $f_{l_c}(l_c)$ the pdf of the Euclidean distance between any two adjacent but disconnected vehicles. It is straightforward that, for $l_c>r_0$

$$f_{l_c}(l_c)=\frac{f_l(l_c)}{1-\int_0^{r_0}f_l(l)dl}=\rho e^{-\rho(l_c-r_0)}. \tag{13}$$

V. Catch-Up Process for a Gaussian Speed Distribution

In this section, we provide detailed analytical results on the catch-up process under the Gaussian speed distributions, which is a commonly used assumption for the VANETs on a highway [3], [9], [15]. The procedure of the analysis is the same as that introduced in the previous section, except for some adjustments to obtain a simpler result.

For a zero-mean Gaussian speed distribution with standard deviation σ, the pdf of the vehicle speed is

$$f_v(v)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-v^2}{2\sigma^2}\right). \tag{14}$$

At the end of the first time slot, i.e., $t=\tau$, it is straightforward to show that

$$p(x,\tau)=\frac{1}{\sigma\tau\sqrt{2\pi}}\exp\left(\frac{-x^2}{2(\sigma\tau)^2}\right). \tag{15}$$

Furthermore, because the convolution of two Gaussian functions is another Gaussian function [21], using (1), we can obtain

$$p(y,i\tau)=\frac{1}{\sigma_i\sqrt{2\pi}}\exp\left(\frac{-y^2}{2\sigma_i^2}\right) \tag{16}$$

where $\sigma_i^2=i(\sigma\tau)^2$.

Next, we consider the situation in which vehicles are allowed to change speed at different time instants. Without loss of generality, consider that a vehicle changes its speed at time τ_0 for the first time since time 0, where τ_0 is uniformly distributed in $(0,\tau]$. Therefore, for $t=i\tau$, (1) becomes

$$\begin{aligned}p(x,t)&=p(x,i\tau)\\&=p(x,\tau_0)*\overbrace{(p*p*\cdots*p)(x,\tau)}^{(\mathrm{i}-1)-\text{fold convolution}}*p(x,\tau-\tau_0)\\&=\frac{1}{\sigma_i\sqrt{2\pi}}\exp\left(\frac{-x^2}{2\sigma_i^2}\right)\end{aligned} \tag{17}$$

where $\sigma_i^2 = (i-1)\sigma^2\tau^2 + \sigma^2\tau_0^2 + \sigma^2(\tau-\tau_0)^2 = i\sigma^2\tau^2 + 2\sigma^2\tau_0^2 - 2\sigma^2\tau\tau_0$.

Compared with the result using a synchronized mobility model [i.e., (16)], the additional terms are $2\sigma^2\tau_0^2 - 2\sigma^2\tau\tau_0$. To simplify the analysis, we ignore the additional terms and consider the synchronized mobility model only. The error caused by ignoring these additional terms and assuming a synchronized mobility model is given by $\int_0^\tau (2\sigma^2\tau_0^2 - 2\sigma^2\tau\tau_0)(1/\tau)d\tau_0 = -\sigma^2\tau^2/3$. It is obvious that the error is small, compared with the dominant term $i\sigma^2\tau^2$, particularly when i is large. Furthermore, the accuracy of this approximation is verified in Section VIII-C.

A. Catch-Up Delay in a Basic Catch-Up Process

In this section, we temporarily ignore the possibility of overtaking, i.e., we consider *a basic catch-up process* involving only the vehicle, which is the head at time 0, catching up with the vehicle, which is the tail at time 0. We have the following lemma for the basic catch-up process:

Lemma 2: In a basic catch-up process where vehicle speed follows a zero-mean Gaussian distribution with standard deviation σ, the probability that the reduction in the distance between the head and the tail is z *for the first time* during the ith time slot is

$$\xi(z,i) = \frac{z}{2i\tau\sigma\sqrt{\pi i}} \exp\left(-\frac{z^2}{4i\tau^2\sigma^2}\right). \tag{18}$$

Proof: Because the Gaussian speed distribution is symmetric with respect to the mean, for the displacement of the tail, we have $p_g(\tilde{y},t) = p(\tilde{y},t) = p(-\tilde{y},t)$. Therefore, using (6), it can be shown that

$$\begin{aligned} p_H(z,i\tau) &= \int_{-\infty}^{\infty} p(y,i\tau)p_g(z-y,i\tau)dy \\ &= (p*p)(z,i\tau) = \frac{1}{\tilde{\sigma}_i\sqrt{2\pi}} \exp\left(\frac{-z^2}{2\tilde{\sigma}_i^2}\right) \end{aligned} \tag{19}$$

where $\tilde{\sigma}_i^2 = 2i(\sigma\tau)^2$. Therefore

$$h(z,i) = \tau p_H(z,i\tau) = \frac{\tau}{\tilde{\sigma}_i\sqrt{2\pi}} \exp\left(\frac{-z^2}{2\tilde{\sigma}_i^2}\right). \tag{20}$$

Inspired by [20, 6.4], $h(z,i)$ in (20) can be rewritten as the following to calculate the Z-transform:

$$h(z,i) = \frac{\tau}{2\pi} \int_{-\infty}^{\infty} \exp\left(-jz\alpha - \frac{\tilde{\sigma}_i^2}{2}\alpha^2\right) d\alpha \tag{21}$$

where j denotes $\sqrt{-1}$.

Then, perform the Z-transform on (21) with regard to i, i.e.,

$$\begin{aligned} (\mathcal{Z}h)(z,s) &= \sum_{i=1}^{\infty} e^{-si} h(z,i) \\ &= \frac{\tau}{2\pi} \int_{-\infty}^{\infty} \exp(-jz\alpha) \sum_{i=1}^{\infty} \exp(-si) \exp\left(-\frac{\tilde{\sigma}_i^2}{2}\alpha^2\right) d\alpha. \end{aligned} \tag{22}$$

With $\tilde{\sigma}_i^2 = 2i(\sigma\tau)^2$ holds.

$$\begin{aligned} (\mathcal{Z}h)(z,s) &= \frac{\tau}{2\pi} \int_{-\infty}^{\infty} \exp(-jz\alpha) \\ &\quad \times \sum_{i=1}^{\infty} \exp(-si) \exp\left(-i(\sigma\tau)^2\alpha^2\right) d\alpha \\ &= \frac{\tau}{2\pi} \int_{-\infty}^{\infty} \exp(-jz\alpha)(s+\sigma^2\tau^2\alpha^2)^{-1} d\alpha \\ &= \frac{\tau}{2} \frac{\exp\left(-z\sqrt{s/(\sigma^2\tau^2)}\right)}{\sqrt{s\sigma^2\tau^2}}. \end{aligned} \tag{23}$$

Then, according to (11), we have

$$(\mathcal{Z}\xi)(z,s) = \frac{(\mathcal{Z}h)(z',s)}{(\mathcal{Z}h)(z'-z,s)} \tag{24}$$

$$= \frac{\exp\left(-z'\sqrt{s/(\sigma^2\tau^2)}\right)}{\exp\left(-(z'-z)\sqrt{s/(\sigma^2\tau^2)}\right)} = \exp\left(-z\sqrt{s/(\sigma^2\tau^2)}\right). \tag{25}$$

Finally, using the inverse Z-transform, it can be obtained that

$$\xi(z,i) = \frac{z}{2i\tau\sigma\sqrt{\pi i}} \exp\left(-\frac{z^2}{4i\tau^2\sigma^2}\right). \tag{26}$$

■

We say that one vehicle *catches up* with another vehicle iff the Euclidean distance between them reduces to the radio range r_0. Then, using Lemma 2, one can readily obtain the following result:

Theorem 1: Consider two vehicles separated by Euclidean distance l_c at time 0. The probability that one vehicle catches up with the other for the first time in the ith time slot is $(l_c - r_0/2i\tau\sigma\sqrt{\pi i}) \exp(-(l_c-r_0)^2/4i\tau^2\sigma^2)$.

Proof: It is straightforward that the reduction in distance is $z = l_c - r_0$. Then, using Lemma 2, the theorem is readily proved. ■

B. Catch-Up Delay With Overtaking Permitted

In the previous section, the possibility of overtaking is not included in the calculation of the first passage probability, to obtain a closed-form result in (26). In this section, the possibility of overtaking is considered to provide a more accurate result on the catch-up delay. Recall that H_m is the mth vehicle to the left of the head H_0 at time 0 and that $P_{m'}$ is the m'th vehicle to the right of the tail P_0 at time 0. Note that all the vehicles H_m ($P_{m'}$) for m, $m' \in [1,\infty)$ can possibly overtake the head H_0 (tail P_0).

Lemma 3: Denote by $q_{mm'}(i|l_c)$ the probability that H_m catches up with $P_{m'}$ $(m, m' \in [0,\infty))$ for the first time in the ith time slot, in a catch-up process with gap l_c. Then, using Lemma 2, we have

$$q_{mm'}(i|l_c) \approx \xi(l_c - r_0 + m/\rho + m'/\rho, i). \tag{27}$$

Proof: Recall that w_m is the distance between H_m and H_0 at time 0, which follows an exponential distribution. Therefore, the expected distance between H_m and H_0 at time 0 is $\int_0^\infty w_m f_{w_m}(w_m)dw_m = m/\rho$, where $f_{w_m}(w_m)$ is given by (4). Similarly, the expected distance between $P_{m'}$ and P_0 at time 0 is m'/ρ. Then, in a catch-up process with gap l_c, the expected distance between H_m and $P_{m'}$ at time 0 is $l_c + m/\rho + m'/\rho$. In order for H_m to catch up with $P_{m'}$, the reduction in distance should be $z = l_c - r_0 + m/\rho + m'/\rho$. ■

Remark 1: In Lemma 3, only the mean value of the distance between vehicles is required. This provides us with the flexibility to extend the analysis from the Poisson distribution model to another vehicle distribution model, i.e., we only need to replace m/ρ in (27) by the corresponding average intervehicle distance if a different vehicle distribution model is used. The rest of the analysis on the catch-up process does not depend on the particular vehicle distribution model being used. However, the accuracy of using mean value approximation for another vehicle distribution needs to be validated.

Denote by $H(i|l_c)$ the probability that none of the $H_m - P_{m'}$ pairs $(m, m' \in [0, \infty))$ catches up in the ith time slot, in a catch-up process with gap l_c. Due to the independence of the movements of vehicles, we have

$$H(i|l_c) = \prod_{m,m' \in [0,\infty)} (1 - q_{mm'}(i|l_c)) \tag{28}$$

where $q_{mm'}(i|l_c)$ is given by Lemma 3. During numerical evaluation, finite values of m, m' can provide fairly accurate results, which are discussed later in Section VIII-A.

Denote by $h(i|l_c)$ the probability that at least one pair of $H_m - P_{m'}$ catches up in the ith time slot and none of them has caught up before the ith time slot, in a catch-up process with gap l_c. It is straightforward that

$$h(i_c|l_c) = (1 - H(i_c|l_c)) \prod_{i=1}^{i_c - 1} H(i|l_c). \tag{29}$$

Finally, the expected delay for a catch-up process with gap l_c is

$$E[t_c|l_c] = \sum_{i=1}^{\infty} i\tau h(i|l_c). \tag{30}$$

VI. Analysis on the Forwarding Process

In a forwarding process, the packet is forwarded in a multihop manner between vehicles inside a cluster. We start with the analysis on the distribution of the length of the cluster.

A. Cluster Length

Define the cluster length x_0 to be the diameter of a cluster, which is the Euclidean distance between the vehicles at the two ends of a cluster. Define $f_{x_0}(x_0)$ to be the pdf of the cluster length, which can be studied as the pdf of the busy period in queuing theory. In previous research, only numerical solutions [3] or approximate results [5] were provided. In this section, we provide a closed-form formula for the pdf of the cluster length using a different method inspired by the study on the connectivity of random interval graph [22] and theory of coverage processes [23].

Theorem 2: In a VANET where the spatial distribution of the vehicles follows a homogeneous Poisson process with intensity ρ, the pdf of the cluster length is

$$f_{x_0}(x_0) = \frac{\rho}{(e^{\rho r_0} - 1)} \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \frac{(-\rho(x_0 - mr_0))^{m-1}}{-m!} \times (\rho(x_0 - mr_0) + m)\, e^{-\rho m r_0} \tag{31}$$

where m is an integer, and $\lfloor . \rfloor$ is the floor function.

The proof is shown in the Appendix.

Theorem 2 shows a closed-form formula for the pdf of the cluster length, which is essential for analytical study on the performance of VANETs.

B. Hop Count Statistics in a Cluster

To study the IPS in the forwarding process, the number of hops between the leftmost vehicle and the rightmost vehicle in the cluster needs to be calculated. Two vehicles are said to be k hops apart if the shortest path between them, which is measured by the number of hops, is k. Define $\phi_k(x_0)$ to be the probability that two vehicles separated by Euclidean distance x_0 are k hops apart. It is assumed that the positions of the vehicles do not change during the forwarding process since the forwarding delay is relatively small, which is also confirmed in Fig. 7. Therefore, the probability $\phi_k(x_0)$ can be calculated by the result introduced in [24] for a static 1-D multi-hop network.

Define $\Pr_s(x_0)$ to be the probability of successful transmissions between any pair of vehicles separated by Euclidean distance x_0. An end-to-end packet transmission is successful if a packet can reach the destination by any number of hops. Therefore, $\Pr_s(x_0) = \sum_{k=1}^{\infty} \phi_k(x_0)$.

Define $\phi_{ks}(x_0)$ to be the conditional probability that a packet reaches its destination at the k_sth hop, conditioned on the transmission being successful and the Euclidean distance between source and destination being x_0. It is trivial to see that $\phi_k(x_0) = \phi_{ks}(x_0) \Pr_s(x_0)$. Therefore, the expected number of hops between two vehicles separated by distance x_0, given that they are connected, is

$$E_s[k_s|x_0] = \frac{\sum_{k=1}^{\infty} k\phi_k(x_0)}{\Pr_s(x_0)}. \tag{32}$$

Define the *forwarding delay* to be the time required for a packet to be forwarded from the leftmost vehicle to the rightmost vehicle in the cluster, which is $t_1 - t_0$ in the example shown in Fig. 1. Then, the expected forwarding delay in a cluster with length x_0 is $E[t_f|x_0] = \beta E_s[k_s|x_0]$.

Remark 2: The hop count statistics for a 1-D network with inhomogeneous Poisson distribution of nodes is studied in [25], which provides us with the required methodology for extending our analysis on the forwarding process from the Poisson distribution model to another vehicle distribution model.

VII. Information Propagation Speed

The entire information propagation process can be considered as a renewal reward process [26], where each cycle consists of a catch-up process, followed by a forwarding process, and the reward is the information propagation distance during each cycle. As mentioned in Section III, $E[v]$ is the constant component of the vehicle speed. It can be shown that [4], [5]

$$E[v_{ip}] \approx \frac{\text{expected length of one cycle}}{\text{expected time duration of one cycle}} + E[v] = \frac{\int_{r_0}^{\infty} l_c f_{l_c}(l_c) dl_c + \int_0^{\infty} x_0 f_{x_0}(x_0) dx_0}{\int_{r_0}^{\infty} E[t_c|l_c] f_{l_c}(l_c) dl_c + \beta + \frac{1}{1-p_c} \int_0^{\infty} E[t_f|x_0] f_{x_0}(x_0) dx_0} + E[v] \quad (33)$$

where $p_c = 2W_{\min}N_b/(W_{\min}+1)^2 + 2W_{\min}N_b$ is the probability of collision given in [27], $W_{\min}$ is the minimum contention window size, and $N_b = \rho 2\pi r_0$ is the average node degree. Packet collision can be shown to have negative impact on the forwarding process, i.e., reducing the IPS, when the vehicle density is high. To illustrate this effect, we conducted more simulations using the parameters shown in [28], i.e., $W_{\min} = 32$.

VIII. Simulation Results

In this section, we report on simulations to validate the accuracy of the analytical results for the catch-up process. The simulations are conducted in a VANET simulator written in C++. Each point shown in the figures is the average value from 2000 simulations. The confidence interval is too small to be distinguishable and, hence, is ignored in the following plots. The radio range is $r_0 = 250$ m [5]. The mobility parameters are $E[v] = 25$ m/s and $\sigma = 7.5$ m/s [15]. To distinguish the impact on the IPS of packet collision and other parameters, we let $p_c = 0$, except in Fig. 11.

A. Catch-Up Process

As mentioned earlier, we use $\tau = 1$ s, 5 s in this paper. Only the results for $\tau = 5$ s are shown in this section since the results for $\tau = 1$ s have a similar (and slightly better) accuracy. The traffic density is $\lambda = 0.3$ veh/s. It follows that the spatial distribution of the vehicles follows a homogeneous Poisson process with intensity $\rho = \int_0^{\infty}(f_v(v)/v)dv = 0.012$ veh/m, which is a low traffic density, resulting in a large number of catch-up processes. The results for other densities are quite similar and, hence, are not shown in this section.

Fig. 5(a) shows the probability that a catch-up process with gap $l_c = 400$ finishes within time t. It can be seen that, when $m = m' = 4$, the analytical result gives a good approximation. Moreover, considering more vehicles in the overtake process, e.g., $m = 6$ or $m = 8$, has marginal impact on the results. This is because, as the distance between vehicles increases, the probability of overtaking rapidly decreases, which can be seen in Fig. 5(b). Fig. 5(b) shows the simulation result of the probability that a randomly chosen vehicle overtakes another vehicle within time $t = i\tau$, where their distance is z at time 0. Because

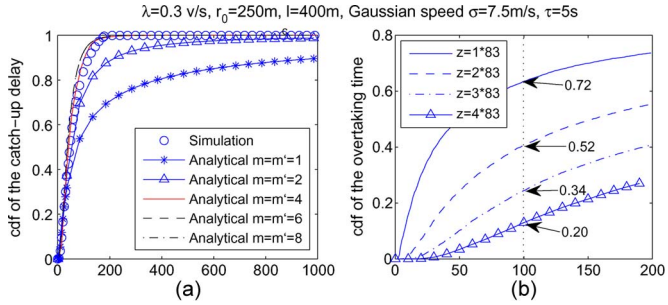


Fig. 5. (a) CDF of the catch-up delay for a catch-up process with gap $l_c = 400$. (b) Simulation results on the probability that a randomly chosen vehicle overtakes another vehicle within time t, where their initial Euclidean distance is $z = m \times 83$ at time 0.

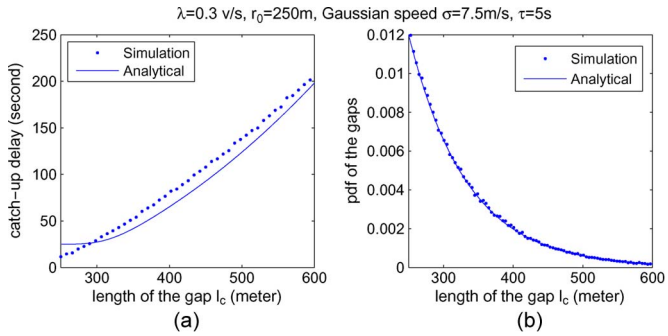


Fig. 6. (a) Expected catch-up delay. (b) PDF of the length of the gap (l_c).

the average intervehicle distance is $1/\rho = 1/0.012 \approx 83$, the curve of $z = m \times 83$ in Fig. 5(b) approximately illustrates the probability that vehicle H_m overtakes H_0 before time t. As can be seen in the figure, the probability that H_4 overtakes H_0 within 100 s while none of H_1, H_2, and H_3 overtakes H_0 within 100 s is approximately given by $0.2(1-0.72)(1-0.52)(1-0.34) = 0.01774$. It can be understood that the probability that H_0 is overtaken by another vehicle, e.g., H_5, $H_6 \ldots$, is very small. Therefore, considering $m = m' = 4$ can provide a good approximation.

Fig. 6(a) shows the expected catch-up delay for a catch-up process with gap l_c. It can be seen that the analytical result, which considers $m = m' = 4$, provides a good approximation. The discrepancy between the simulation result and the analytical result is caused by the approximations used during the analysis. Specifically, the first passage analysis is only applied to the analysis of the catch-up process between a pair of vehicles, which are the head and the tail at the start of the catch-up process. However, the first passage analysis does not consider the possibility that the head (the tail) may be overtaken by other vehicles during the catch-up process. Furthermore, Fig. 6(b) also verifies that the intervehicle distance, under our network model and the Gaussian speed distribution, still follows an exponential distribution with $\rho = 0.012$. This property is also expected to hold in some other speed distributions, which is an issue that is left as future work.

B. Forwarding Process

In addition to the simulation settings introduced earlier, the per-hop delay is $\beta = 4$ ms [5]. Fig. 7(a) shows the expected

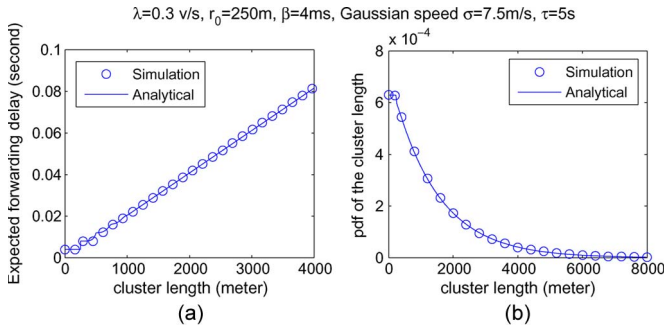


Fig. 7. Expected forwarding delay and the pdf of the cluster length. (a) Expected forwarding delay. (b) PDF of the cluster length.

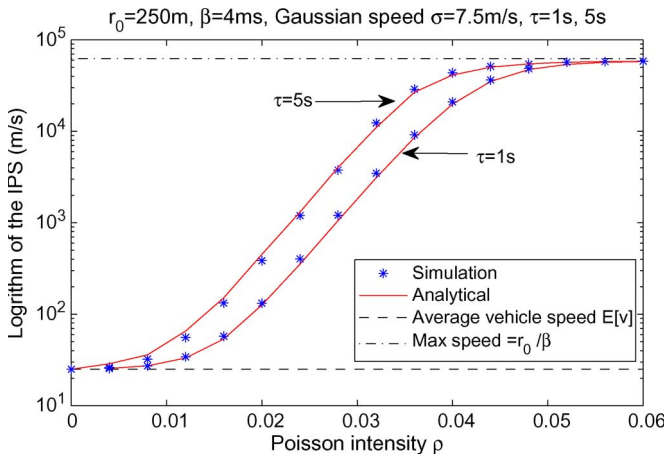


Fig. 8. Expected IPS for $\tau = 1$ s, 5 s.

forwarding delay in a cluster with a given length. Fig. 7(b) shows the pdf of the cluster length. It can be seen that the analytical results match the simulation results very well. The results for other values of the parameters have a similar accuracy and are thus omitted. Furthermore, it is interesting to note that, in Fig. 7(b), the pdf of the cluster length is a constant for $x_0 \in [0, 250]$. This is because, within the radio range $(r_0 = 250)$ of the first vehicle, $\Pr(\mathcal{E}_4) = 1$ [(36)], and the cluster length is x_0 if and only if there is a vehicle in $[x_0, x_0 + dx_0)$, and there is no vehicle in $[x_0 + dx, x_0 + r_0)$. It follows that the pdf of the cluster length is a constant for $x_0 \in [0, r_0]$.

C. IPS

In addition to the simulation settings introduced earlier, the Poisson arrival rate λ is varied from 0 to 1.5 veh/s. With $E[v] = 25$, the spatial distribution of the vehicles follows a homogeneous Poisson process with intensity ρ ranging from 0 to 0.06. For completeness of the plot, $\rho = 0$ is included, which means that there is only one vehicle on the road. Therefore, the average number of neighbors (average node degree) varies from 0 to 30, which represents a large range of traffic densities.

Fig. 8 shows the expected IPS for $\tau = 1$ s, 5 s. It can be seen that, when the vehicle density is low, the IPS is determined by vehicle speeds because there is little packet forwarding in the network. When the vehicle density increases, small clusters are formed, and the IPS is determined by the catch-up delay, which is further determined by the mobility of the vehicles. It can be

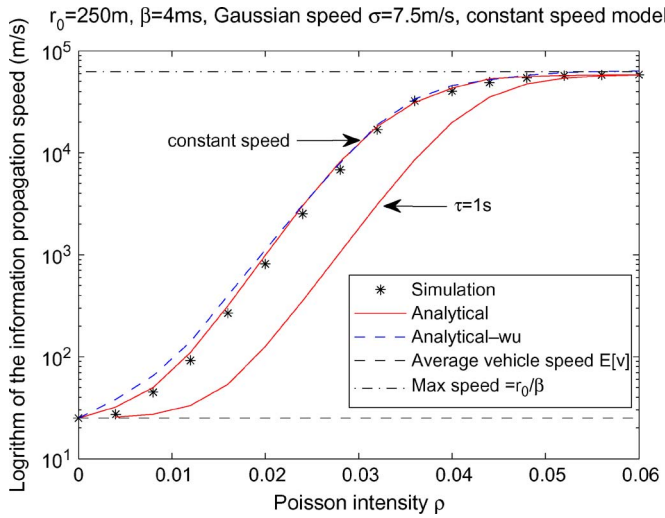


Fig. 9. Expected IPS under the constant speed model, together with the curve for $\tau = 1$ for comparison. The result *Analytical-wu* is calculated based on [5].

seen that the more frequently the speed changes, the slower the information propagates. This is mainly because changing speed has the potential to interrupt the catch-up process, i.e., during a catch-up process, the tail may speed up and the head may slow down. An intuitive explanation can be provided by considering an extreme case. In the extreme case that the speed-change time interval tends to 0, it can be shown using the central limit theorem that the average vehicular speed in any specified time interval converges to the mean speed $E[v]$. Hence, the network topology becomes static, and the expected IPS equals the mean speed $E[v]$ because there is no catch-up process to bridge the gaps. Finally, as the vehicle density further increases, clusters become larger, and the forwarding process starts to dominate. Therefore, the IPS increases until it reaches the maximum value, which is determined by the per-hop delay in the forwarding process. The maximum IPS is obviously equal to r_0/β, where β is the per-hop delay.

Fig. 9 shows the expected IPS under the constant speed model, i.e., the vehicle speed does not change over time. The result *Analytical-wu* is calculated based on [5] for comparison. The constant speed model is a special case of the mobility model used in this paper, i.e., when $\tau \to \infty$. In Fig. 9, we choose a fairly large value of τ to obtain analytical result under our mobility model and use the result as an approximation of the result under the constant speed model. The result *Analytical-wu*, which does not consider first passage phenomenon, provides a good match with the simulations. This is because the first passage phenomenon does not have a significant impact when the vehicle speed does not change over time. Finally, it can be seen that the constant speed model used in previous research (e.g., [5] and [8]) causes serious overestimation of the IPS by almost an order of magnitude. Therefore, time variation of vehicle speed is an important factor affecting the IPS.

Fig. 10 shows the expected IPS under nonsynchronized mobility models. In the analysis, we choose the synchronized random walk mobility model to study the impact of the time variation of vehicular speed on IPS. To verify the general applicability of the analytical study based on the simplified

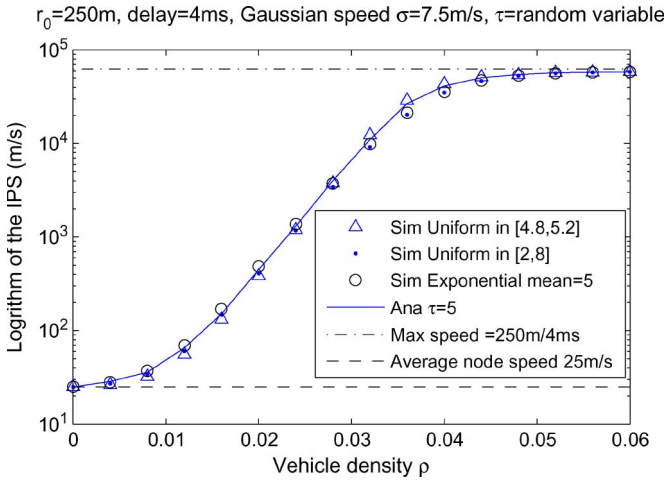


Fig. 10. Expected IPS under nonsynchronized mobility models.

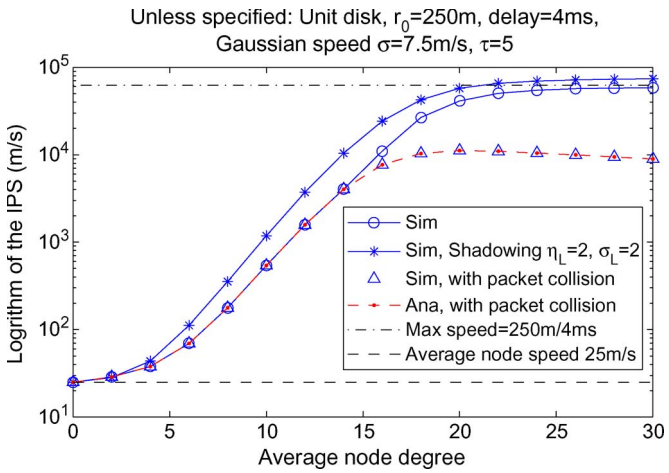


Fig. 11. Expected IPS in a VANET subject to log-normal shadowing and packet collision, where η_L is the path-loss exponent, and σ_L is the standard deviation of the log-normal shadowing. Furthermore, when a head cannot transmit a packet to any uninformed vehicle to the right of itself, the head keeps trying to retransmit the packet after every 0.9-s time delay. The value is chosen according to the real-world measurement that the channel coherent time of a VANET on the freeway is about 0.3–1.5 s [29].

mobility model, more simulations are conducted. As shown in Fig. 10, three different mobility models are evaluated, i.e., the speed-change time interval τ of each vehicle is uniformly selected from [4.8, 5.2] (i.e., about 5) or from [2, 8] (i.e., within a larger range around 5) and for τ following an exponential distribution with mean 5. Under all three mobility models, vehicles change speed at different time instants (nonsynchronized), and the average speed-change time interval is 5 s. It can be seen that the IPSs under nonsynchronized mobility models are very close (almost indistinguishable) to each other, and our analysis using the simplified (synchronized) mobility model provides a good estimation on the IPS.

We consider the unit disk model in the analysis. The unit disk communication model is constructed based on the path-loss attenuation model, which is suitable to model the radio environment in free space without clutters [30]. Therefore, the unit disk model is suitable for the VANET on the freeway. To study the impact of clutters such as road-side buildings, simulation results of the expected IPS under the log-normal shadowing model [30] are shown in Fig. 11. The results are compared under the condition that the average node degrees (i.e., the average number of neighbors per node) under the log-normal model and under the unit disk model are the same. In the log-normal shadowing model, the received signal strength (RSS) attenuation (in decibels) follows a normal distribution with a mean value equal to the RSS under the path-loss attenuation model. This random variation on the RSS attenuation provides a higher chance for a node to find a next-hop neighbor. Hence, even with the same average node degree, the IPS under log-normal shadowing model is faster than that under unit disk model. A similar observation is obtained in the study of network connectivity in [7] and [30, Th. 2.5.2]. Therefore, the IPS under the unit disk model can be considered as a lower bound on the IPS of a VANET in the real world.

In addition, the third and fourth curves in Fig. 11 show the IPS subject to packet collision with collision probability p_c given in Section VII. It can be seen that the packet collision has a significant impact on the IPS, particularly when the vehicle density is high.

IX. Conclusion and Future Work

In this paper, analytical models have been created for the information propagation process in a mobile ad hoc network formed by vehicles moving on a highway. Analytical results have been provided for the expected delay in catch-up process, expected delay in forwarding process, and the distribution of the cluster length. Based on the aforementioned results, the IPS has been derived. It has been shown that various parameters, such as vehicle density and time variation of vehicle speed, can have a significant impact on the IPS. By taking the real-world measurements such as λ, $E[v]$, σ, and τ, our results can provide a quick estimation of the IPS with good accuracy. The research provides useful guidelines on the design of mobile VANETs.

The analysis in this paper is conducted in a 1-D network. A straightforward extension to 2-D networks is to consider a Manhattan model [32], i.e., grid topology. In grid topology, each street (edge) can be treated as a 1-D roadway, which can be studied using the results obtained in this paper. Together with existing results on how a message passes through a road intersection area [33], our results can be adapted for 2-D grid networks. Furthermore, for unconstrained mobility in 2-D or 3-D mobile ad-hoc networks, the methodology developed in this paper for the analysis of the catch-up process and forwarding process can also be applicable.

Appendix
Proof of Theorem 2

Let the origin of the axis be the position of the leftmost vehicle of a cluster. Let N be the random variable representing the number of vehicles in the cluster. The cluster length lies in $[x_0, x_0 + dx_0)$, and there are n vehicles in this cluster if and only if the conditions given here hold.

$\mathcal{E}_1$: There is a vehicle in $[x_0, x_0 + dx_0)$.

$\mathcal{E}_2$: There is no vehicle in $[x_0 + dx_0, x_0 + r_0)$.

$\mathcal{E}_3$: There are $n - 2$ vehicles in $(0, x_0)$.

$\mathcal{E}_4$: The intervehicle distance between any two adjacent vehicles for those n vehicles in $[0, x_0 + dx_0)$ is smaller than or equal to r_0.

Denote by $\Pr(\mathcal{E}_m)$ the probability of event $\mathcal{E}_m$ in the preceding list. Due to the Poisson distribution of vehicles, it is straightforward to show that

$$\Pr(\mathcal{E}_1) = \rho dx_0, \quad \Pr(\mathcal{E}_2) = e^{-\rho r_0} \tag{34}$$

$$\Pr(\mathcal{E}_3) = \frac{(\rho x_0)^{n-2} e^{-\rho x_0}}{(n-2)!}. \tag{35}$$

Furthermore, $\Pr(\mathcal{E}_4)$ can be studied using [22, Lemma 1], which provides result on the connectivity of random interval graph. In [22], vertices are uniformly distributed on a unit interval. Due to the Poisson distribution of the vehicles in our case, given that there are n vehicles in a cluster with length x_0, these vehicles also follow a uniform distribution. Therefore, by scaling the cluster length x_0 to 1 and, consequently, the radio range to r_0/x_0, we have the following equation from Lemma 1 in [22]:

$$\Pr(\mathcal{E}_4) = \sum_{m=0}^{\min\{n-1, \lfloor x_0/r_0 \rfloor\}} \binom{n-1}{m} (-1)^m \left(1 - m\frac{r_0}{x_0}\right)^{n-2} \tag{36}$$

where m is an integer, and $\lfloor . \rfloor$ is the floor function. For the convenience of the following calculation, let $\binom{n-1}{m} = 0$ for $m > n-1$. Thus, the preceding summation is from $m = 0$ to $\lfloor x_0/r_0 \rfloor$.

Events $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_3$ are independent of each other. Event $\mathcal{E}_4$, which is conditioned on event $\mathcal{E}_3$, is independent of events $\mathcal{E}_1$ and $\mathcal{E}_2$. Define $f(x_0, N = n)$ to be the probability that the cluster length lies in $[x_0, x_0 + dx_0)$, and there are n vehicles in this cluster. It is evident that $f(x_0, N = n) = \Pr(\mathcal{E}_1)\Pr(\mathcal{E}_2)\Pr(\mathcal{E}_3)\Pr(\mathcal{E}_4)$. Next, we derive $f_{x_0}(x_0)$ using $f(x_0, N = n)$.

If the cluster only consists of one vehicle, then the cluster length is 0, and the probability of this event is $e^{-\rho r_0}$. If the cluster consists of more than one vehicle, then the pdf of cluster length x_0 is

$$\begin{aligned} f_{x_0}(x_0) &= \frac{\sum_{n=2}^{\infty} f(x_0, N = n)}{\Pr(N \geq 2)} \\ &= \sum_{n=2}^{\infty} \frac{\rho e^{-\rho r_0} (\rho x_0)^{n-2} e^{-\rho x_0}}{(n-2)!(1 - e^{-\rho r_0})} \\ &\quad \times \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \binom{n-1}{m} (-1)^m \left(1 - m\frac{r_0}{x_0}\right)^{n-2} \\ &= \frac{\rho e^{-\rho r_0} e^{-\rho x_0}}{(1 - e^{-\rho r_0})} \sum_{n=2}^{\infty} \frac{(\rho x_0)^{n-2}}{(n-2)!} \\ &\quad \times \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \binom{n-1}{m} (-1)^m \left(1 - m\frac{r_0}{x_0}\right)^{n-2}. \end{aligned} \tag{37}$$

Using $\binom{n}{m} = n!/m!(n-m)!$, the summation terms in (37) can be simplified as

$$\begin{aligned} &\sum_{n=2}^{\infty} \frac{(\rho x_0)^{n-2}}{(n-2)!} \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \binom{n-1}{m} (-1)^m \left(1 - m\frac{r_0}{x_0}\right)^{n-2} \\ &= \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \frac{(-1)^m}{m!} \sum_{n=2}^{\infty} \frac{(n-1)!}{(n-2)!} \frac{\left(\rho x_0 \left(1 - m\frac{r_0}{x_0}\right)\right)^{n-2}}{(n-m-1)!} \\ &= \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \frac{(-1)^m}{m!} \sum_{n=2}^{\infty} \frac{(n-1)\left(\rho(x_0 - mr_0)\right)^{n-2}}{(n-m-1)!} \\ &= \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \frac{(-1)^m}{m!} \left(\rho(x_0 - mr_0)\right)^{m-1} \\ &\quad \times \sum_{n=2}^{\infty} \frac{(n-1)\left(\rho(x_0 - mr_0)\right)^{n-m-1}}{(n-m-1)!} \\ &= \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \frac{(-1)^m}{m!} \left(\rho(x_0 - mr_0)\right)^{m-1} \\ &\quad \times \sum_{n=m+1}^{\infty} \frac{(n-1)\left(\rho(x_0 - mr_0)\right)^{n-m-1}}{(n-m-1)!} \end{aligned} \tag{38}$$

$$\begin{aligned} &= \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \frac{(-1)^m}{m!} \left(\rho(x_0 - mr_0)\right)^{m-1} \\ &\quad \times \sum_{\alpha=0}^{\infty} \frac{(\alpha + m)\left(\rho(x_0 - mr_0)\right)^{\alpha}}{\alpha!}. \end{aligned} \tag{39}$$

Using $xe^x = \sum_{\alpha=0}^{\infty} (\alpha x^{\alpha}/\alpha!)$, (39) can be written as

$$\begin{aligned} &\sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \frac{(-1)^m}{m!} \left(\rho(x_0 - mr_0)\right)^{m-1} \\ &\qquad \times \left(\rho(x_0 - mr_0) + m\right) e^{\rho(x_0 - mr_0)}. \end{aligned} \tag{40}$$

Substitute the preceding equation into (37), one can obtain

$$\begin{aligned} f_{x_0}(x_0) &= \frac{\rho e^{-\rho r_0} e^{-\rho x_0}}{(1 - e^{-\rho r_0})} \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \frac{(-1)^m}{m!} \left(\rho(x_0 - mr_0)\right)^{m-1} \\ &\quad \times \left(\rho(x_0 - mr_0) + m\right) e^{\rho(x_0 - mr_0)} \\ &= \frac{\rho e^{-\rho r_0}}{(1 - e^{-\rho r_0})} \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \frac{(-1)^m}{m!} \left(\rho(x_0 - mr_0)\right)^{m-1} \\ &\quad \times \left(\rho(x_0 - mr_0) + m\right) e^{-\rho m r_0} \\ &= \frac{\rho}{(e^{\rho r_0} - 1)} \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \frac{(-1)^m}{m!} \left(\rho(x_0 - mr_0)\right)^{m-1} \\ &\quad \times \left(\rho(x_0 - mr_0) + m\right) e^{-\rho m r_0} \\ &= \frac{\rho}{(e^{\rho r_0} - 1)} \sum_{m=0}^{\lfloor x_0/r_0 \rfloor} \frac{\left(-\rho(x_0 - mr_0)\right)^{m-1}}{-m!} \\ &\quad \times \left(\rho(x_0 - mr_0) + m\right) e^{-\rho m r_0}. \end{aligned} \tag{41}$$

■

REFERENCES

[1] H. Hartenstein and K. P. Laberteaux, "A tutorial survey on vehicular ad hoc networks," *IEEE Commun. Mag.*, vol. 46, no. 6, pp. 164–171, Jun. 2008.

[2] I. Berger, *Standards for Car Talk*, The Institute, Mar. 2007.

[3] S. Yousefi, E. Altman, R. El-Azouzi, and M. Fathy, "Analytical model for connectivity in vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 6, pp. 3341–3356, Nov. 2008.

[4] A. Agarwal, D. Starobinski, and T. D. Little, "Exploiting downstream mobility to achieve fast upstream message propagation in vehicular ad hoc networks," in *Proc. Mobile Netw. Veh. Environ.*, 2007, pp. 13–18.

[5] H. Wu, R. M. Fujimoto, G. F. Riley, and M. Hunter, "Spatial propagation of information in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 420–431, Jan. 2009.

[6] H. Wu, J. Lee, M. Hunter, R. Fujimoto, R. L. Guensler, and J. Ko, "Efficiency of simulated vehicle-to-vehicle message propagation in Atlanta, Georgia, I-75 corridor, *Transp. Res. Rec.: J. Transp. Res. Board*, vol. 1910, pp. 82–89, 2005.

[7] R. Hekmat and P. V. Mieghem, "Connectivity in wireless ad-hoc networks with a log-normal radio model," *Mobile Netw. Appl.*, vol. 11, no. 3, pp. 351–360, Jun. 2006.

[8] A. Agarwal, D. Starobinski, and T. D. Little, "Analytical model for message propagation in delay tolerant vehicular ad hoc networks," in *Proc. IEEE Veh. Technol. Conf.*, 2008, pp. 3067–3071.

[9] W. Leutzbach, *Introduction to the Theory of Traffic Flow*. New York: Springer-Verlag, 1988.

[10] P. Jacquet, B. Mans, and G. Rodolakis, "Information propagation speed in mobile and delay tolerant networks," in *Proc. INFOCOM*, 2009, pp. 244–252.

[11] R. Fracchia and M. Meo, "Analysis and design of warning delivery service in intervehicular networks," *IEEE Trans. on Mobile Comput.*, vol. 7, no. 7, pp. 832–845, Jul. 2008.

[12] D. Camara, C. Bonnet, and F. Filali, "Propagation of public safety warning messages: A delay tolerant network approach," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2010, pp. 1–6.

[13] Q. Xu, R. Sengupta, T. Mak, and J. Ko, "Vehicle-to-vehicle safety messaging in DSRC," in *Proc. 1st ACM Int. Workshop Veh. ad hoc Netw.*, 2004, pp. 19–28.

[14] *Traffic software integrated system—Corridor simulation*, 2009, U. S. Dep. Transp.

[15] M. Rudack, M. Meincke, and M. Lott, "On the dynamics of ad hoc networks for inter vehicle communications (IVC)," in *Proc. ICWN*, Las Vegas, NV, 2002.

[16] C. Bettstetter, "Mobility modeling in wireless networks: Categorization, smooth movement, and border effects," *SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 5, no. 3, pp. 55–66, Jul. 2001.

[17] Carchip fleet pro, 2010. [Online]. Available: http://www.carchip.com/Product_Docs

[18] R. Nelson, *Probability, Stochastic Processes, and Queueing Theory: The Mathematics of Computer Performance Modeling*. New York: Springer-Verlag, 1995.

[19] S. Redner, *A Guide to First-Passage Processes*. Cambridge, U.K.: Cambridge Univ. Press, 2001.

[20] E. Montroll and B. West, "On an enriched collection of stochastic processes," in *Fluctuation Phenomena, Studies in Statistical Mechanics*. Amsterdam, The Netherlands: North-Holland, 1979, pp. 61–175.

[21] C. M. Grinstead and J. L. Snell, *Introduction to Probability*. Providence, RI: AMS, 1997.

[22] E. Godehardt and J. Jaworski, "On the connectivity of a random interval graph," *Random Struct. Algorithms*, vol. 9, no. 1/2, pp. 137–161, Sep. 1996.

[23] P. Hall, *Introduction to the Theory of Coverage Processes*. Hoboken, NJ: Wiley, 1988.

[24] S. Dulman, M. Rossi, P. Havinga, and M. Zorzi, "On the hop count statistics for randomly deployed wireless sensor networks," *Int. J. Sensor Netw.*, vol. 1, no. 1/2, pp. 89–102, Jan. 2006.

[25] A. Zanella, G. Pierobon, and S. Merlin, "On the limiting performance of broadcast algorithms over unidimensional ad-hoc radio networks," in *Proc. WPMC*, Abano Terme, Italy, 2004.

[26] S. M. Ross, *Introduction to Probability Models*. New York: Academic, 2007.

[27] J. Li and C. Chigan, "Delay-aware transmission range control for vanets," in *Proc. IEEE GLOBECOM*, 2010, pp. 1–6.

[28] M. M. Carvalho and J. J. Garcia-Luna-Aceves, "Delay analysis of IEEE 802.11 in single-hop networks," in *Proc. IEEE Int. Conf. Netw. Protocols*, 2003, pp. 146–155.

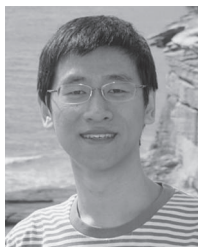
[29] L. Cheng, B. E. Henty, D. D. Stancil, F. Bai, and P. Mudalige, "Mobile vehicle-to-vehicle narrow-band channel measurement and characterization of the 5.9 GHz dedicated short range communication (DSRC) frequency band," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 8, pp. 1501–1516, 2007.

[30] T. S. Rappaport, *Wireless Communications: Principles and Practice,* 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2001.

[31] M. Franceschetti and R. Meester, *Random Networks for Communication: From Statistical Physics to Information Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

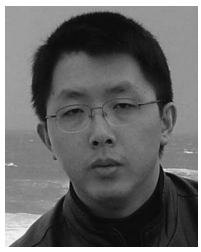
[32] F. Bai, N. Sadagopan, and A. Helmy, "Important: A framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks," in *Proc. IEEE INFOCOM*, 2003, pp. 825–835.

[33] J. Lee, G.-L. Park, I.-H. Shin, and M.-J. Kang, "Design of intersection switches for the vehicular network," in *Proc. Manage. Enabling Future Internet Changing Bus. New Comput. Serv.*, 2009, vol. 5787, pp. 523–526.



Zijie Zhang (S'11) received the B.Eng. degree in electronic and communications engineering from the University of Hong Kong, Kowloon, Hong Kong, in 2007 and is currently working toward the Ph.D. degree with the School of Electrical and Information Engineering, University of Sydney, Sydney, Australia.

His research interests include wireless multihop networks, vehicular ad-hoc networks, mobile ad-hoc networks, and graph theory.



Guoqiang Mao (SM'08) received the Ph.D. degree in telecommunications engineering from Edith Cowan University, Perth, Australia, in 2002.

In December 2002, he joined the School of Electrical and Information Engineering, University of Sydney, Sydney, Australia, where he is currently a Senior Lecturer. His research interests include wireless localization techniques, wireless multihop networks, graph theory and its applications in networking, and network performance analysis.

Brian D. O. Anderson (M'66–SM'74–F'75–LF'07) was born in Sydney, Australia. He received the B.Sc. degree in mathematics and the B.E. degree in electrical engineering from Sydney University, in 1962 and 1964, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1966.

He is currently a Distinguished Professor with the Research School of Information Sciences and Engineering, Australian National University, Canberra, Australia, and a Distinguished Researcher with National ICT Australia, Canberra. His current research interests are distributed control, sensor networks, and econometric modeling.

Dr. Anderson is a Fellow of the Australian Academy of Science, the Australian Academy of Technological Sciences and Engineering, and the Royal Society, and a foreign associate of the National Academy of Engineering. He was the recipient of the IEEE Control Systems Award in 1997, the 2001 IEEE James H Mulligan, Jr. Education Medal, and the Bode Prize from the IEEE Control System Society in 1992, as well as several IEEE best paper prizes.

On the Information Propagation Process in Multi-lane Vehicular Ad-hoc Networks

Zijie Zhang*, Guoqiang Mao*‡ and Brian D. O. Anderson†‡
*School of Electrical and Information Engineering, University of Sydney, Australia
†Research School of Engineering, Australian National University, Australia
‡National ICT Australia (NICTA), Australia
Email: {zijie.zhang, guoqiang.mao}@sydney.edu.au, brian.anderson@anu.edu.au

***Abstract*—This paper studies the information propagation process in a 1D mobile ad-hoc network formed by vehicles traveling on a highway. We consider that vehicles can be divided into traffic streams; vehicles in the same traffic stream have the same speed distribution, while the speed distributions of vehicles in different traffic streams are different. Analytical formulas are derived for the fundamental properties of the information propagation process as well as the information propagation speed. Using the formulas, one can straightforwardly study the impact on the information propagation speed of various parameters such as radio range, vehicular traffic density, vehicular speed distribution and the time variation of vehicular speed.**

I. Introduction

This paper studies the *information propagation speed (IPS)*, which is the expected propagation speed for a piece of information to be broadcast along the road in a vehicular ad-hoc network (VANET). The IPS is an important performance metric for many VANET applications, especially for the safety messaging applications [1].

It has been shown [2], [3] that a VANET is usually partitioned into a number of clusters, where a *cluster* is a maximal set of vehicles in which every pair of vehicles is connected by at least one multi-hop path. Due to the mobility of vehicles, the clusters are splitting and merging over time. Therefore, information propagation in a VANET is typically based on a store-and-forward scheme. Considering the example illustrated in Fig. 1, a piece of information starts to propagate from the origin toward the positive direction of the axis at time t_0. The vehicles that have received this piece of information are referred to as the *informed vehicles*, where other vehicles are *uninformed*. The multi-hop forwarding of the message within a cluster, which begins at t_0 and ends at t_1, is called a *forwarding process*. In a forwarding process the information propagation speed is determined by the per-hop delay and the length of the cluster. The *per-hop delay* β is the time required for a vehicle to receive and process a message before it is available for further retransmission [4]. The value of β reflecting the typical technology is 4ms [4]. We show that the per-hop delay has a significant impact on the IPS.

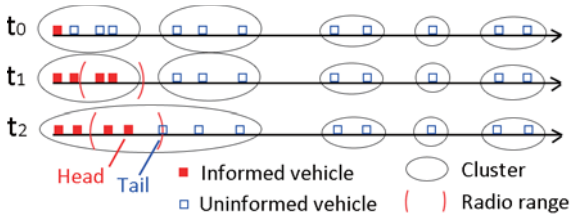


Fig. 1. Illustration of the topology of a VANET at different time instants. The positive direction of the axis is the direction of information propagation.

Define the *head* at time t as the informed vehicle with the largest coordinate at time t. Define the *tail* at time t as the uninformed vehicle with the smallest coordinate at time t. Two vehicles can directly communicate with each other if and only if their Euclidean distance is not larger than the radio range r_0, i.e. we adopt the *unit disk model*. As shown in Fig. 1, at time t_1 the tail is outside the radio range of the head. Then a *catch-up process* begins, during which the informed vehicles hold the information until the head catches up the tail (at time t_2).

Recent research has shown that downstream traffic (a set of vehicles traveling in the opposite direction of information propagation) can be exploited to improve the IPS [2], [5], [6]. Further, real world measurements show that vehicles traveling in different lanes (e.g. bus lane or heavy truck lane) have different speed distributions [7]. In view of these observations, we consider multiple traffic streams, where a *traffic stream* is a set of vehicles following the same speed distribution. A traffic stream does not have to consist of vehicles traveling in the same lane. Traffic streams also represent different types of vehicles (e.g. sports cars or heavy trucks). Interesting results are obtained by allowing vehicles in different traffic streams to have different speed distributions.

Main contributions of this paper are: firstly, an analytical model for the information propagation process considering multiple traffic streams is provided. It is shown that a small difference in the average vehicular speeds between traffic streams can result in a significant increase of the IPS. Secondly, time variation of vehicular speed is considered in the analysis, which results in interesting conclusions, e.g. the IPSs are similar for information propagating in both positive and

This research is partially supported by ARC Discovery project DP110100538. This material is based on research partially sponsored by the Air Force Research Laboratory, under agreement number FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

negative directions of a road, which is different from previous studies (e.g. [5], [6]) considering time-invariant vehicular speed only. Based on the analysis of the catch-up process and forwarding process, analytical results for the IPS are derived, which are validated using simulations. This paper provides useful guidelines on the design of a multi-lane VANET.

The rest of this paper is organized as follows: Section II reviews related work. Section III introduces the mobility model and network model. The analysis on the catch-up process is given in Section IV. The forwarding process is studied in Section V, followed by the results of the IPS. Section VI validates the analysis using simulations. Section VII concludes this paper and discusses future work.

II. Related work

In [2], Agarwal et al. studied the IPS in a 1D VANET where vehicles are Poissonly distributed and move at the same speed, which is time-invariant, but either in the positive (upstream) or negative (downstream) direction of the axis. The upstream and downstream traffic have the same density. This work was later extended in [5] to allow the upstream and the downstream traffic to have different densities. The authors derived upper and lower bounds for the IPS, which provided a hint on the impact of vehicle density on the IPS. Recent research of Baccelli et al. [6] provided analytical results on the IPS under the same setting as that in [5]. Note that the analyses in [2], [5], [6] were all based on the simplifying assumption that vehicles in each traffic stream travel at the same speed, which is time-invariant. Evidently, in such model, catch-up can only occur via the vehicles travelling in the opposite direction.

In [3], Wu et al. considered a 1D VANET where vehicles are Poissonly distributed and the vehicle speeds are uniformly distributed within a designated range. Considering one traffic stream and time-invariant vehicular speed, they provided analytical results on the IPS when the vehicle density is either very low or very high. They also provided a numerical method to calculate the IPS in a VANET with two traffic streams. In an earlier work [8], we showed that the time-variation of vehicular speed has a significant impact on the IPS. However, only one traffic stream was considered in [8], i.e. all vehicles travel in the same direction and follow the same speed distribution. This paper considers multiple traffic streams and evaluating the impact of vehicular densities and speed distributions on information propagation in multi-lane VANETs.

III. System Model

Suppose that there are a total of N traffic streams, wherein vehicles are free to change lanes and overtake other vehicles. We adopt the commonly-used traffic model in traffic theory [9], viz. the number of vehicles in the n^{th} traffic stream passing an observation point during any time interval follows a homogeneous Poisson process with intensity λ_n veh/s.

A synchronized random walk mobility model is considered. Specifically, time is divided into time slots with equal length τ. Then, the i^{th} time slot is $t \in ((i-1)\tau, i\tau]$. Each vehicle in the n^{th} ($n \in [1, N]$) traffic stream chooses its speed randomly at the beginning of each time slot, independent of the speeds of other vehicles and its own speed in other time slots, according to a probability density function (pdf) $f_{vn}(v)$. In this paper, we consider the Gaussian speed distribution, i.e. $f_{vn}(v) \sim \mathcal{N}(\mu_n, \sigma_n^2)$, where μ_n (resp. σ_n^2) is the mean speed (resp. variance) in the n^{th} traffic stream. The Gaussian speed distribution is commonly used for modeling the VANETs on a freeway [3], [9], [10].

Under the aforementioned setting, it can be shown that at any time instant the spatial distribution of the vehicles in the n^{th} traffic stream follows a homogeneous Poisson process with intensity $\rho_n = \lambda_n \int_{-\infty}^{\infty} \frac{f_{vn}(v)}{v} dv$ [8]. Then according to the superposition property of the Poisson process, at any time instant the spatial distribution of all the vehicles on the road follows a homogeneous Poisson process with intensity $\rho = \sum_{n=1}^{N} \rho_n$.

IV. Catch-up process

Without loss of generality, it is assumed that the catch-up process starts at time 0. The *displacement* x ($x \in (-\infty, \infty)$) of a vehicle at time t is defined to be the distance between the position of the vehicle at time 0 and its position at time t.

This section is organized as follows: in Lemma 1, we study the movement of a single vehicle. The result on the distance between two vehicles at a given time is summarized in Lemma 2. After considering the overtaking of vehicles in Section IV-C, we obtain the delay for a catch-up process in Section IV-D.

A. Modeling the movement of a single vehicle

Denote by $p_n(x,t)$ the probability that the displacement of a vehicle in the n^{th} traffic stream is x at time t. We have:

Lemma 1: Under the system model introduced in Section III, when $t = i\tau$,

$$p_n(x,t) = p_n(x, i\tau) = \frac{1}{\tilde{\sigma}_i \sqrt{2\pi}} \exp\left(\frac{-(x - \mu_n i\tau)^2}{2\tilde{\sigma}_i^2}\right) \quad (1)$$

where $\tilde{\sigma}_i^2 = i(\sigma_n \tau)^2$.

Proof: According to Gaussian speed distribution model, the pdf of the speed of a vehicle in the n^{th} traffic stream is:

$$f_{vn}(v) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left(\frac{-(v - \mu_n)^2}{2\sigma_n^2}\right) \quad (2)$$

Because the speed does not change during a time slot, it is straightforward that $p_n(x,\tau)$ is also a Gaussian function:

$$p_n(x,\tau) = \frac{1}{\sigma_n \tau \sqrt{2\pi}} \exp\left(\frac{-(x - \mu_n \tau)^2}{2(\sigma_n \tau)^2}\right) \quad (3)$$

Due to the independence of the vehicular speeds in different time slots (hence the displacements), we have for $t = i\tau$:

$$p_n(x,t) = p_n(x, i\tau) = \overbrace{(p_n * p_n * \ldots * p_n)(x,\tau)}^{\text{i-fold convolution}} \quad (4)$$

Then one can obtain Eq. 1 using the property that the convolution of two Gaussian functions is a Gaussian function. ■

B. The distance between a pair of vehicles

As shown in Fig. 2, denote by H_m (resp. P_m) the m^{th} vehicle to the left of the head H_0 (resp. to the right of the tail P_0) at time 0. If H_m happens to be in the n^{th} traffic stream, which happens with a probability $\frac{\rho_n}{\rho}$, then an additional label n (e.g. H_m^n) is used to indicate that the vehicle H_m is in the n^{th} traffic stream. In this subsection, we study the distance between H_m^n and $P_{m'}^{n'}$, where $n, n' \in [1, N]$ and m, m' are non-negative integers.

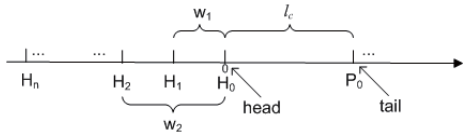


Fig. 2. Illustration of a VANET at the beginning of a catch-up process.

The catch-up process between H_m^n and $P_{m'}^{n'}$ finishes as soon as the distance between them reduces to radio range r_0 *for the first time*. Note that the reduction of distance between H_m^n and $P_{m'}^{n'}$ can reach a given value z at several time instants with non-zero probabilities. We are interested in the probability that the reduction of the distance between H_m^n and $P_{m'}^{n'}$ reaches z *for the first time*, i.e. the first passage time phenomenon, which is essential for the analysis of the catch-up process.

Lemma 2: Denote by $G_{mm'}^{nn'}(z,i)$ the probability that the reduction of the distance between H_m^n and $P_{m'}^{n'}$, with regards to their distance at time 0, reaches z ($z \neq 0$) *for the first time* in the i^{th} ($i \geq 1$) time slot. Then: (a) When $\mu_n = \mu_{n'}$:

$$G_{mm'}^{nn'}(z,i) = \frac{z}{i\sqrt{2\pi i\tau^2(\sigma_n^2+\sigma_{n'}^2)}}\exp(\frac{-z^2}{2i\tau^2(\sigma_n^2+\sigma_{n'}^2)}) \quad (5)$$

(b) When $\mu_n \neq \mu_{n'}$ and $i = 1$:

$$G_{mm'}^{nn'}(z,i) \leq \frac{1}{2}(1-erf(\frac{z-\mu_i}{\sqrt{2\sigma_i^2}})) \quad (6)$$

(c) When $\mu_n \neq \mu_{n'}$ and $i \geq 2$:

$$G_{mm'}^{nn'}(z,i) \leq \frac{1}{4}(1+erf(\frac{z-\mu_{i-1}}{\sqrt{2\sigma_{i-1}^2}}))(1-erf(\frac{z-\mu_i}{\sqrt{2\sigma_i^2}})) \quad (7)$$

where $\mu_i = \mu_n i\tau - \mu_{n'} i\tau$, $\sigma_i^2 = (\sigma_n^2+\sigma_{n'}^2)\tau^2 i$.

Proof: Denote by $g_{mm'}^{nn'}(z,t)$ the probability that the reduction of the distance between H_m^n and $P_{m'}^{n'}$ is z at time t. When $t = i\tau$,

$$\begin{aligned} g_{mm'}^{nn'}(z,t) &= \int_{-\infty}^{\infty} p_n(x,t)p_{n'}(x-z,t)dx & (8)\\ &= \int_{-\infty}^{\infty} p_n(x,t)p_{n'}(z-x,t)dx & (9)\\ &= (p_n * p_{n'})(x,t) & (10)\\ &= \frac{1}{\sqrt{2\pi\sigma_i^2}}\exp(\frac{-(z-\mu_t)^2}{2\sigma_t^2}) & (11)\end{aligned}$$

where $\mu_t = (\mu_n - \mu_{n'})t$, $\sigma_t^2 = (\sigma_n^2+\sigma_{n'}^2)\tau t$ and Eq. 11 is obtained from Eq. 10 using Lemma 1. Note that Eq. 9 is obtained from Eq. 8 by letting the mean speed of $P_{m'}^{n'}$ be 0 instead of $\mu_{n'}$, and consequently the mean speed of H_m^n becomes $\mu_n - \mu_{n'}$. Therefore, $p_{n'}(x,t)$ becomes a zero mean Gaussian function with respect to x hence $p_{n'}(x-z,t) = p_{n'}(z-x,t)$.

When two traffic streams have the same mean speed, i.e. $\mu_n = \mu_{n'}$, we have a special case that $\mu_t = 0$ in Eq. 11. For simplicity, the probability that the reduction of the distance between H_m^n and $P_{m'}^{n'}$ reaches z in the i^{th} time slot, i.e. $\int_{(i-1)\tau}^{i\tau} g_{mm'}^{nn'}(z,t)dt$ is approximated by $\tau g_{mm'}^{nn'}(z,i\tau)$. That is we approximately consider that $g_{mm'}^{nn'}(z,t) = g_{mm'}^{nn'}(z,i\tau)$ for $t \in ((i-1)\tau, i\tau]$. Applying a standard procedure [8] to determine the first passage probability $G_{mm'}^{nn'}(z,i)$, we have:

$$\tau g_{mm'}^{nn'}(z',i'\tau) = \tau\sum_{i=1}^{i'} G_{mm'}^{nn'}(z,i)g_{mm'}^{nn'}(z'-z,(i'-i)\tau) \quad (12)$$

Solving the equation by the Z-transform as introduced in our previous work [8], one can obtain that for $\mu_n = \mu_{n'}$:

$$G_{mm'}^{nn'}(z,i) = \frac{z}{i\sqrt{2\pi i\tau^2(\sigma_n^2+\sigma_{n'}^2)}}\exp(\frac{-z^2}{2i\tau^2(\sigma_n^2+\sigma_{n'}^2)}) \quad (13)$$

When $\mu_n \neq \mu_{n'}$, we use a method different from the complicated first passage analysis to calculate $G_{mm'}^{nn'}(z,i)$ from $g_{mm'}^{nn'}(z,t)$. Define $\mathcal{A}$ to be the event that the reduction of distance between H_m^n and $P_{m'}^{n'}$ is smaller than z at time $(i-1)\tau$. Define $\mathcal{B}$ to be the event that the reduction of distance between H_m^n and $P_{m'}^{n'}$ is larger than z at time $i\tau$. Then for $\mu_n \neq \mu_{n'}$ and $i \geq 2$:

$$\begin{aligned} G_{mm'}^{nn'}(z,i) &\leq \Pr(\mathcal{A})\Pr(\mathcal{B}|\mathcal{A}) \leq \Pr(\mathcal{A})\Pr(\mathcal{B}) & (14)\\ &= \int_0^z g_{mm'}^{nn'}(z_0,(i-1)\tau)dz_0\int_z^{\infty} g_{mm'}^{nn'}(z_0,i\tau)dz_0 & (15)\\ &= \frac{1}{2}(1+erf(\frac{z-\mu_{i-1}}{\sqrt{2\sigma_{i-1}^2}}))\frac{1}{2}(1-erf(\frac{z-\mu_i}{\sqrt{2\sigma_i^2}})) & (16)\end{aligned}$$

where $\mu_i = \mu_n i\tau - \mu_{n'} i\tau$, $\sigma_i^2 = (\sigma_n^2+\sigma_{n'}^2)\tau^2 i$. The first inequality is due to the fact that we do not consider the first passage phenomenon. Hence $\Pr(\mathcal{A})\Pr(\mathcal{B}|\mathcal{A})$ can be larger than the probability that the reduction of distance between H_m^n and $P_{m'}^{n'}$ reaches z *for the first time* in the i^{th} time slot. This bound is fairly tight in the case $\mu_n \neq \mu_{n'}$ that we considered here. Because when $\mu_n \neq \mu_{n'}$, $g_{mm'}^{nn'}(z,t)$ becomes a Gaussian function with a non-zero mean according to Eq. 11. It follows that the mean reduction of distance between H_m^n and $P_{m'}^{n'}$ becomes a decreasing (when $\mu_n < \mu_{n'}$) or increasing (when $\mu_n > \mu_{n'}$) function of t. Therefore the first passage phenomenon is less notable when $\mu_n \neq \mu_{n'}$. Secondly, the bound $\Pr(\mathcal{B}|\mathcal{A}) \leq \Pr(\mathcal{B})$ in Eq. 14 is due to the observation that $\mathcal{A}$ and $\mathcal{B}$ are negatively correlated, i.e. given that the reduction of distance is less than z at time $(i-1)\tau$, the reduction of distance is less likely to be larger than z at time $i\tau$. Further, this correlation appears to have limited impact on the final result under the settings introduced in Section VI.

For $\mu_n \neq \mu_{n'}$ and $i = 1$, $\Pr(\mathcal{A}) = 1$. Therefore:

$$G_{mm'}^{nn'}(z,i) \leq \Pr(\mathcal{B}) = \frac{1}{2}(1-erf(\frac{z-\mu_i}{\sqrt{2\sigma_i^2}})) \quad (17)$$

■

C. The catch-up process between H_m and $P_{m'}$

A catch-up process, where the distance between the head H_0 and tail P_0 is l_c at time 0, is referred to as a *catch-up process with gap* l_c. Note that the head at time t is not necessarily the head vehicle at time 0 (H_0) because H_0 may be overtaken by another informed vehicle during time $(0, t]$. We first consider the catch-up process between a pair of vehicles H_m-$P_{m'}$.

Denote by $q_{mm'}(i|l_c)$ the probability that H_m catches up $P_{m'}$ for the first time in the i^{th} time slot, in a catch-up process with gap l_c. Recall that with probability $\frac{\rho_n}{\rho}$ (resp. $\frac{\rho_{n'}}{\rho}$) that H_m (resp. $P_{m'}$) belongs to the n^{th} (resp. n'^{th}) traffic stream. It can be shown that

$$q_{mm'}(i|l_c) = \sum_{n,n' \in [1,N]} G^{nn'}_{mm'}(z,i) \frac{\rho_n \rho_{n'}}{\rho^2} \tag{18}$$

where $z = l_c - r_0 + w_m + w_{m'}$ and w_m (resp. $w_{m'}$) is the expected distance between H_m and H_0 (resp. $P_{m'}$ and P_0) at time 0 as shown in Fig. 2. Due to the Poisson distribution of vehicles, the inter-vehicle distance follows an exponential distribution with mean $1/\rho$. It follows that $w_m = m/\rho$, $w_{m'} = m'/\rho$. Therefore $z = l_c - r_0 + m/\rho + m'/\rho$.

D. Delay of a catch-up process

Define the *catch-up delay* t_c to be the time taken from the beginning of a catch-up process till the time when the head and tail move into the radio range of each other for the first time, i.e. $t_2 - t_1$ in Fig. 1. Denote by $H(i|l_c)$ the probability that none of the H_m-$P_{m'}$ pairs catches up in the i^{th} time slot, in a catch-up process with gap l_c. Due to the independence between the movements of vehicles, we have:

$$H(i|l_c) = \prod_{m,m' \in [0,\infty)} (1 - q_{mm'}(i|l_c)) \tag{19}$$

where $q_{mm'}(i|l_c)$ is given by Eq. 18.

Denote by $h(i|l_c)$ the probability that at least one pair of H_m-$P_{m'}$ catches up in the i^{th} time slot *and* none of them has caught up before the i^{th} time slot, in a catch-up process with gap l_c. Assume that the catch-up event in the i^{th} time slot is independent of that in the j^{th} time slot for $0 < j < i$, which is an accurate approximation when the duration of a time slot is large, e.g. $\tau = 1$s or 5s as shown in Section VI. Then:

$$h(i_c|l_c) = (1 - H(i_c|l_c)) \prod_{i=1}^{i_c - 1} H(i|l_c) \tag{20}$$

Finally the expected delay for a catch-up process with gap l_c is $E[t_c|l_c] = \sum_{i=1}^{\infty} i\tau h(i|l_c)$

V. Forwarding process and IPS

Define *forwarding delay* as the time required for a packet to be forwarded from the leftmost vehicle to the rightmost vehicle in a cluster. Assume that a cluster does not become disconnected during the forwarding process since the per-hop delay is relatively small (e.g. 4ms [4]). Then, the expected forwarding delay in a cluster with length x_0 is $E[t_f|x_0] = \beta E[k|x_0]$, where $E[k|x_0]$ (given in [8]) is the expected number of hops between two vehicles separated by x_0.

The entire information propagation process can be considered as a renewal reward process where each cycle consists of a catch-up process followed by a forwarding process and the reward is the information propagation distance during each cycle. Therefore, the expected IPS ($E[v_{ip}]$) is [3]:

$$\begin{aligned} E[v_{ip}] &\approx \frac{\text{expected length of one cycle}}{\text{expected time duration of one cycle}} + \mu_{max} \\ &= \frac{\int_{r_0}^{\infty} l_c f_{l_c}(l_c) dl_c + \int_0^{\infty} x_0 f_{x_0}(x_0) dx_0}{\int_{r_0}^{\infty} E[t_c|l_c] f_{l_c}(l_c) dl_c + \beta + \int_0^{\infty} E[t_f|x_0] f_{x_0}(x_0) dx_0} \\ &\quad + \mu_{max} \end{aligned} \tag{21}$$

where $\mu_{max} = \max_{n \in N}\{\mu_n\}$ is the maximum average speed among traffic streams and $f_{l_c}(l_c) = \rho e^{-\rho(l_c - r_0)}$ is the pdf of distance between two adjacent but disconnected vehicles [8].

VI. Simulation results

Simulations are conducted in a VANET simulator written in C++. Each point shown in the figures is the average value from 2000 simulations. The confidence interval is too small to be distinguishable and hence is ignored in the following plots. The radio range is $r_0 = 250m$ [3]. The typical values of the mean and standard deviation of vehicular speed are 25m/s and 7.5m/s [10]. These mobility parameters, i.e. μ_n, σ_n and τ, are taken from practical measurements, where the usual record time intervals for a vehicle speed monitor are $\tau = 1s, 5s$ [11]. The traffic density is varied so that the density ρ varies from 0 to 0.06 veh/m. For completeness of the plot, $\rho = 0$ is included which means there is only one vehicle in each traffic stream.

Fig. 3 and Fig. 4 show the expected IPS in a VANET. Firstly it can be seen that analytical results have a good match with simulation results. Secondly, when the vehicular density is low, the IPS is determined by vehicular speeds (whose mean value is 25m/s) because there is little packet forwarding. Thirdly, when the vehicular density increases, small clusters are formed and the IPS increases. As the vehicular density further increases, clusters become larger and the forwarding process starts to dominate. Therefore, the IPS increases until it reaches the maximum value r_0/β, where β is the per-hop delay. When the vehicular density is moderate, the IPS is determined by the catch-up delay, which is further determined by the mobility of vehicles, as shown in our previous work [8].

Moreover, vehicular speed distribution also has a significant impact on the catch-up process hence the IPS. Fig. 3 shows the IPS in a VANET with two traffic streams with equal vehicular density but different vehicular speed distributions. It can be seen that our analytical result has a better match with the simulation results than that in [6], which does not consider the catch-up process studied in Lemma 2 of this paper. An interesting observation is that the IPS increases when the average speed of the vehicles in one of the traffic streams is reduced from 25m/s to 0m/s. Further, an even faster IPS is observed when the average speed of the vehicles in one of the traffic streams is -25m/s, i.e. two traffic streams head in opposite directions. The reason behind this interesting observation is that a larger relative speed between vehicles

results in faster catch-up processes hence a faster IPS. This can also be seen from the analytical results of $G_{mm'}^{nn'}(z,i)$, e.g. applying the fact that the error function is an increasing function to Eq. 6. This observation tells us that making use of the vehicles in the negative traffic stream can increase the IPS. Moreover, some (stationary) roadside units without (expensive) wired connections can also significantly increase the IPS in a VANET.

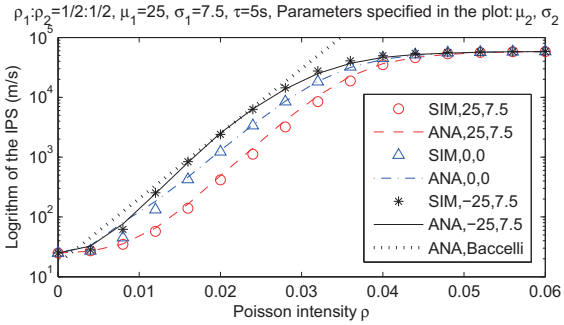


Fig. 3. The expected IPS in a VANET with different vehicular speeds in two traffic streams ($N = 2$). The plot $Baccelli$ shows the IPS derived in [6] for a VANET with vehicular speed 25 and -25 in two streams respectively.

Fig. 4 shows the expected IPS in a VANET with different vehicular densities in two traffic streams. Firstly, it can be seen that an uneven distribution of vehicular densities between traffic streams (e.g. $\rho_1 = \rho/10, \rho_2 = 9\rho/10$) results in a slower IPS, compared with the IPS in a VANET with evenly distributed traffic densities (e.g. $\rho_1 = \rho_2 = \rho/2$). This is because an uneven distribution of vehicular densities between traffic streams results in a smaller number of catch-ups between vehicles in different traffic streams (as manifested by the factor $\frac{\rho_n \rho_{n'}}{\rho^2}$ in Eq. 18), hence less improvement on IPS is provided by the large relative vehicular speed, compared with a VANET with evenly distributed vehicular densities between traffic streams. This situation can be observed on freeways connecting the business district and residential district. Densities of the traffic streams in opposite directions can vary greatly depending on time-of-day when people going to work or coming back home. On the other hand, it can be seen that a small amount (e.g. $\rho/10$) of vehicular traffic in an opposite direction can still significantly increase the IPS. Therefore, it is meaningful to exploit the traffic streams with different mean speeds (e.g. heavy trucks or buses).

Moreover, it can be seen in Fig. 4 that the last four curves are very close to each other, which suggests that given a configuration of the vehicular densities and speed distributions in a VANET, the IPSs are similar for the information propagating in both positive and negative directions. This conclusion is different from previous ones that were based on the time-invariant speed model [5], [6]. Our results suggest that a similar IPS is achievable in both directions of a two-way communication despite the uneven vehicular density between traffic streams.

VII. Conclusion and Future Work

In this paper, analytical results are provided for the information propagation process in a mobile ad-hoc network

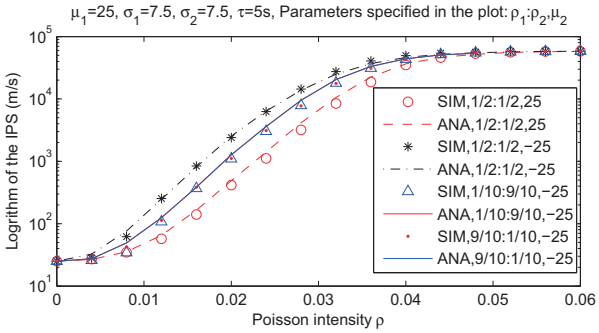


Fig. 4. The expected IPS in a VANET with different vehicular densities in two traffic streams. The first four curves are kept from Fig. 3 for comparison.

formed by vehicles moving on a freeway. It is considered that vehicles can be divided into multiple traffic streams. Vehicles in the same traffic stream have the same speed distribution while the speed distributions of vehicles in different traffic streams are different. We found that the IPS can be boosted significantly by exploiting the existence of even a small number of vehicles traveling with a different mean speed, such as vehicles traveling in the opposite direction or heavy trucks moving in the same direction but with a slower speed. On the other hand, the IPSs are similar for the information propagating toward both positive and negative directions in the case of uneven traffic densities among various traffic streams. By using parameters extracted from real world measurements for λ, μ, σ and τ, our results can provide a quick estimate of the IPS with a good accuracy. The analysis can be further extended to 2D topologies in the future. Further, we are going to consider the impact of channel randomness on the IPS.

References

[1] D. Camara, C. Bonnet, and F. Filali, "Propagation of public safety warning messages: a delay tolerant network approach," in *Proceedings IEEE WCNC*, 2010.

[2] A. Agarwal, D. Starobinski, and T. D. Little, "Analytical model for message propagation in delay tolerant vehicular ad hoc networks," in *IEEE Vehicular Technology Conference*, 2008, pp. 3067–3071.

[3] H. Wu, R. M. Fujimoto, G. F. Riley, and M. Hunter, "Spatial propagation of information in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 420–431, 2009.

[4] H. Wu, J. Lee, M. Hunter, R. Fujimoto, R. L. Guensler, and J. Ko, "Efficiency of simulated vehicle-to-vehicle message propagation in Atlanta, Georgia, I-75 corridor," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1910, pp. 82–89, 2005.

[5] A. Agarwal and T. D. Little, "Impact of asymmetric traffic densities on delay tolerant vehicular ad hoc networks," in *IEEE Vehicular Networking Conference (VNC)*, 2009, pp. 1 – 8.

[6] E. Baccelli, P. Jacquet, B. Mans, and G. Rodolakis, "Information propagation speed in bidirectional vehicular delay tolerant networks," in *Proceedings IEEE INFOCOM*, 2011, pp. 436–440.

[7] G. Yan, N. Mitton, and X. Li, "Reliable routing in vehicular ad hoc networks," in *The 7th International Workshop on Wireless Ad hoc and Sensor Networking*, Genoa, Italie, 2010.

[8] Z. Zhang, G. Mao, and B. D. Anderson, "On the information propagation speed in mobile vehicular ad hoc networks," in *Proceedings IEEE GLOBECOM*, 2010, pp. 1–5.

[9] W. Leutzbach, *Introduction to the Theory of Traffic Flow*. Springer-Verlag, 1988.

[10] M. Rudack, M. Meincke, and M. Lott, "On the dynamics of ad hoc networks for inter vehicle communications (IVC)," in *Proceedings ICWN*, 2002.

[11] "Carchip fleet pro," 2010. [Online]. Available: http://www.carchip.com

On Information Dissemination in Infrastructure-based Mobile Ad-hoc Networks

Zijie Zhang*, Guoqiang Mao*‡ and Brian D. O. Anderson†‡
*School of Electrical and Information Engineering, University of Sydney, Australia
†Research School of Information Sciences and Engineering, Australian National University, Australia
‡National ICT Australia (NICTA), Australia
Email: {zijie.zhang, guoqiang.mao}@sydney.edu.au, brian.anderson@anu.edu.au

***Abstract*—In this paper, we consider 2D wireless multi-hop networks with mobile nodes randomly distributed on a torus, and a small number of base stations (infrastructure nodes) deterministically placed in the same area. Mobile nodes move following a random walk mobility model. A piece of information is broadcast from the base stations at the same time in a multi-hop manner using a Susceptible-Infectious-Recovered (SIR) epidemic routing algorithm. A distinguishing feature of the SIR algorithm, which leverages the mobility of mobile users, is that a relay node carries a piece of information for a pre-determined amount of time and forwards it at any available opportunity during that time. We provide analytical results for the percolation probability and for the expected fraction of nodes that receive the information when the information dissemination process stops. Further, we study the time delay of the information dissemination process. The accuracy of the analytical results is verified using simulations.**

***Index Terms*—mobile ad-hoc networks, infrastructure, percolation, random walk, epidemic routing**

I. Introduction

In this paper, we consider an infrastructure-based mobile ad-hoc network (MANET) with a large number of mobile nodes (MNs) moving following a random walk mobility model and a small number of base stations / infrastructure nodes (INs). Examples of such infrastructure-based MANETs include wildlife tracking sensor networks, vehicular ad-hoc networks and mobile social networks [1]. Taking the mobile social network as an example, the information (news or advertisement) dissemination relies not only on the mobile broadband connection between mobile users and the base stations, but also on the local ad-hoc connections between mobile users that emerge as mobile users move and meet each other. This ad-hoc method of information dissemination can reduce the resource usage of base stations and lower the cost for service providers and users by utilizing cheaper radio resources, such as Bluetooth or cognitive radio, and physical mobility of the mobile users, for information dissemination. Moreover, in the case of natural disasters causing damage to or loss of a base station, ad-hoc communications between mobile devices may become crucial for survivors. In this paper, we investigate the information dissemination process in infrastructure-based MANETs.

The topology of a MANET often resembles the topology of a human network, in the sense that the mobility of nodes in a MANET is not only similar to, but often governed by, the movements of their human owners. In view of this, epidemic routing algorithms [2] have been proposed as a fast and reliable approach to disseminate information in MANETs. An epidemic routing algorithm adopts the so-called *store-carry-forward* paradigm, where a mobile node stores and carries its received information and then forwards the information to its neighbors when direct links to them emerge. As a consequence of the store-carry-forward paradigm, the information is forwarded from a source to a destination using a journey instead of a path, where a *journey* is an alternation of packet transmissions and carriages, that connects a source to a destination [3]. The epidemic routing algorithm leverages the mobility of nodes and efficiently disseminates information in MANETs [2], [4].

In this paper, we study the information dissemination process in a 2D infrastructure-based MANET, where a piece of information is broadcast from the infrastructure nodes to the mobile nodes in a multi-hop manner using a Susceptible-Infectious-Recovered (SIR) epidemic routing algorithm, which is introduced in detail in Section III. An asymptotic upper bound is derived for the percolation probability, viz. the probability that a non-vanishingly-small fraction of MNs receive the information as the number of MNs approaches infinity. Then we study the expected fraction of MNs that receive the information when the information dissemination process stops. Further, we derive a lower bound on the time delay of the information dissemination process. Finally, the accuracy of the analytical results is verified using simulations.

The rest of this paper is organized as follows: Section II reviews the related work. Section III introduces the network model and the epidemic routing algorithm considered in this paper. The analysis on the information dissemination process is given in Section IV. Section V validates the analysis using simulations. Finally Section VI concludes this paper and proposes possible future work.

This research is partially supported by ARC Discovery project DP110100538. This material is based on research partially sponsored by the Air Force Research Laboratory, under agreement number FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

II. Related work

Some related work treating infrastructure-based ad-hoc networks can be found in the literature in the context of static networks. A static network is a special case of a MANET when the MNs are stationary, in which case the MNs are referred to as the ordinary nodes (ONs). Alternatively, a static network can be considered to be a snapshot of a MANET. In [5], Bermudez and Wicker considered a network where ONs are uniformly distributed in a unit square, with one IN at each corner of the square. They empirically investigated the fraction of ONs that are connected to at least one IN by Monte Carlo simulations. In [6], Zhang et al. considered a network with ONs Poissonly distributed and INs deterministically placed in a given 2D area. They studied the fraction of ONs which are connected to at least one IN in at most k hops. In [7], Dousse et al. considered a network where ONs are Poissonly distributed in $\Re^2$ and INs are placed on a grid. Using simulations, they showed that an increase in the number of INs provides little improvement on the probability that an arbitrary ON is connected to at least one IN, when the density of ONs is either very high or very low. However, the benefit of INs for the connectivity of networks with an intermediate density of ONs was not investigated. Note that all the above results are for static networks.

In view of the special characteristics of dynamic networks, epidemic routing algorithms [2], [4] have been proposed to disseminate information in MANETs. In [8], Zyba et al. studied the performance of a Susceptible-Infectious (SI) epidemic routing algorithm using real world mobility traces. By separating users into two behavioral classes, they obtained the successive meeting time of two nodes and the fraction of nodes that receive the information broadcast from an arbitrary source. In [9], Zhang et al. studied analytically the performance of SIR epidemic routing in a MANET where nodes are Poissonly distributed on a torus and move following a random direction model. They studied the fraction of MNs that receive the information broadcast from an arbitrary MN. However, the aforementioned studies considered only ad-hoc connection between MNs in a MANET without infrastructure support.

Time delay is an important metric for both end-to-end and broadcast information disseminations. In [10], Groenevelt et al. studied the end-to-end information dissemination in a MANET using an unrestricted multi-copy protocol like the SI epidemic routing algorithm. Using the approximation that the successive meeting time of two nodes follows an exponential distribution, they showed that the end-to-end delay between an arbitrary pair of nodes converges to $\frac{\log N}{0.57721N}$ asymptotically as the number of nodes $N \to \infty$. In [11], Zhang et al. studied the performance of several epidemic routing algorithms in MANETs assuming that the successive meeting time of two nodes follows an exponential distribution with rate β. By solving ordinary differential equations, they showed a similar result to that in [10], i.e. the expected end-to-end delay between an arbitrary pair of nodes is $\frac{\log N}{\beta N}$ using the SI epidemic routing algorithm, where N is the number of nodes in the network. In [12], Zhou et al. studied the performance of SI epidemic routing in a MANET with N mobile nodes and M base stations. At each time slot, a single pair of transmitter-receiver nodes is uniformly and randomly selected among $N+M$ nodes in the network. It follows that the rate of transmission is once per time slot, irrespective of the number of MNs in the network. Modeling the information propagation process by a Markov chain, they found that the expected time required to deliver a message from an arbitrary MN to at least one of the base stations is $\Theta(N \log N)$. The analysis in this paper is not based on the assumption on the inter-meeting time or the rate of transmission. Further, we show the impact of various parameters, i.e. node density, radio range and a quality-of-service metric (fraction of nodes that receive the information), on the delay of the information dissemination in an infrastructure-based MANET.

III. System model

A. Network model

Consider a MANET where mobile nodes (MNs) are randomly and independently distributed on a torus $(0, L]^2$ [13] following a homogeneous Poisson point process with intensity λ. It follows that the expected number of MNs in the network is $N = \lambda L^2$. Further, M infrastructure nodes (INs) are deterministically placed in the same area, in a way that they are not clustered together, as shown in Fig. 1. An example is to place the INs in a way that the associated Voronoi cells of these INs have equal size. We consider that the number of INs is much less than the number of MNs, i.e. $M \ll N$, which reflects the fact that INs are usually more expensive. Two nodes are directly connected iff (if and only if) their Euclidean distance is smaller than or equal to the radio range r_0, viz. we adopt the unit disk communication model. The use of torus allows a node located near the boundary to have the same number of connections probabilistically as a node located near the center. Therefore, the torus topology is a simplification that helps to obtain analytical results, because there is no need to consider the boundary effect, and it is commonly used in this area [10], [13].

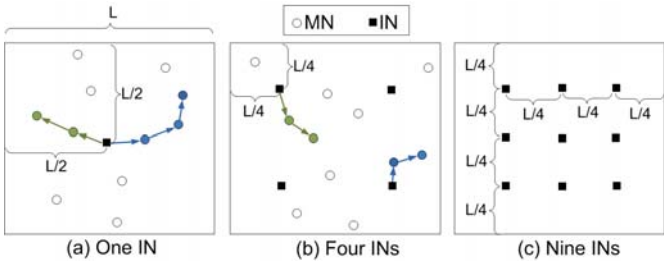


Fig. 1. Examples of the deployment of one, four or nine infrastructure nodes.

Further, MNs move according to the random walk model (RWM) [14], following which each MN moves at a constant speed V in a particular direction for a time duration following an exponential distribution with mean ζ seconds. The direction is uniformly chosen in $[0, 2\pi)$, independent of the direction of other MNs and the node's direction in other time intervals.

Note that when $\zeta \to \infty$, i.e. the direction does not change over time, and our analysis therefore provides the results for a MANET with the random direction model (RDM) [14]; when $\zeta \to 0$ the mobility model is like Brownian motion; when $V \to 0$, our analysis provides the results for a static network.

It is worth noting that under the aforementioned node distribution model and mobility models, the spatial distribution of the MNs is stationary and always follows a homogeneous Poisson point process with intensity λ at any time instant [15].

B. Susceptible-Infectious-Recovered (SIR) routing algorithm

Consider a basic stochastic SIR epidemic routing algorithm [9], [16], where a piece of information is broadcast from the INs to the MNs at time $t = 0$. We consider that the INs are connected to each other by a backbone network so that INs can broadcast the same piece of information at the same time.

By analogy to the way a disease spreads in a human network, a MN can be in any of three states S, I, R: A MN that has never received the information from any of the INs is in the state of susceptible (S), in which the MN can accept incoming transmissions if such opportunity arises. A susceptible node transits into the state of infected and infectious (I) immediately after it has received a copy of the information. A node in state I keeps transmitting the information, i.e. remains infectious, for a certain time period τ, which is referred to as the *active period*. Note that τ is a pre-determined value which is the same for all nodes. After the active period, the node enters into the state of recovered and immune (R). A node in state R stops transmitting the information to other nodes and will ignore all future transmissions of the same information from other nodes. An animation of the information dissemination process in a MANET is available on [17].

Note that the INs only transmit a piece of information for a time period of length τ, after which the information propagates in the network using the epidemic routing algorithm. The information dissemination naturally stops (i.e. reaches the *steady state*) when there is no infectious node in the network. We study the fraction of informed MNs when the information dissemination process has reached the steady state, where the *informed MNs* are MNs that have received the information.

The network introduced in this subsection is denoted by $\mathcal{G}(L, \lambda, M, V, \zeta, \tau)$ hereafter. Further, we consider that the network is deployed on a sufficiently large torus (i.e. $L > V\tau$) such that within the active period τ, a node will not be wrapped back to the point where it starts moving at time 0.

IV. Analysis of the information dissemination process

A. The definitive metric

The information dissemination process of a MANET is determined by a number of parameters such as node density, mobility, radio range and the length of active period. In this subsection, a single metric is proposed, which captures the impact of the above parameters.

Definition 1. The *effective node degree* R_0 of an infectious node is the expected number of MNs that have been inside an infectious node's radio range during the infectious node's active period.

Note that R_0 is the same for all MNs because of the stationarity and homogeneity of node distribution on the torus. Further, we do not need to consider the INs in the calculation of R_0, because the INs are the sources of the information and they do not act as a relay which receives an information then re-broadcast it. Next we provide an upper bound on R_0.

Lemma 1. *Consider a network $\mathcal{G}(L, \lambda, M, V, \zeta, \tau)$. The effective node degree R_0 satisfies:*

$$R_0 \leq \frac{8r_0 V \lambda \tau}{\pi} + \pi r_0^2 \lambda \tag{1}$$

Proof: We first study the size of the area covered by the radio range of a typical mobile node (Q) during its active period assuming that its direction does not change. As shown in Fig. 2, the area consists of two parts. The first part is the area swept by a segment of length $2r_0$ perpendicular to the trajectory of the node, which is shown by the line-shaded area and is referred to as the *area swept by the radio range*. The second part is the area of two half-circles, which is shown by the uniformly shaded area and whose size is πr_0^2.

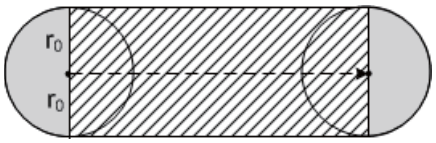


Fig. 2. An illustration of the area (line-shaded and uniformly shaded) covered by the radio range of a node when the node moves along a straight line.

Note that if a MN changes its direction over time, then this will cause overlapping of the trajectory and a reduction in the size of the area swept by the radio range during its active period. However, an analytical upper bound on this area, call it $\tilde{A}(V)$, can be obtained by ignoring the reduction of the area caused by the overlapping of the trajectory. It is straightforward that $\tilde{A}(V) \leq 2r_0 V\tau$.

Denote by Θ the angle measured counterclockwise from the direction of the node Q to the direction of an arbitrary MN. Recall that the direction of an arbitrary MN is randomly and uniformly chosen in $[0, 2\pi)$, independent of the directions of other nodes. It can be shown that Θ is uniformly distributed in $[0, 2\pi)$. Based on the thinning property of a Poisson process, it can be shown that the subset of MNs moving toward directions $\Theta \in (\theta, \theta + d\theta)$ follows a homogeneous Poisson process with intensity $\lambda d\theta$. Further, the relative speed between node Q and the aforementioned subset of MNs is $\sqrt{V^2 + V^2 - 2V^2 \cos\theta} = 2V|\sin\frac{\theta}{2}|$. To facilitate the analysis, we consider a new coordinate system where the origin is set at an arbitrary node moving toward direction θ. In this new coordinate system, the locations of the aforementioned subset of MNs are approximately stationary[1] and the node

[1]The locations of the MNs may have small displacements during time interval τ due to a small direction difference $d\theta$. However the displacements become vanishingly small when $d\theta \to 0$ while V and τ are finite.

Q is moving at speed $2V|\sin\frac{\theta}{2}|$. It follows that in the new coordinate system, the expected size of area swept by the radio range of node Q is $\tilde{A}(2V|\sin\frac{\theta}{2}|)$. Therefore the expected number of MNs moving toward directions $\Theta \in (\theta, \theta + d\theta)$, that have been inside the area swept by the radio range of node Q during its active period, is $\tilde{A}(2V|\sin\frac{\theta}{2}|)d\lambda\theta$.

Considering all mobile nodes in the network, it is straightforward that the expected number of MNs that have been inside the area swept by the radio range of node Q is:

$$\int_0^{2\pi} \tilde{A}(2V|\sin\frac{\theta}{2}|)\frac{1}{2\pi}\lambda d\theta \tag{2}$$

$$\leq \int_0^{2\pi} 2r_0 2V|\sin\frac{\theta}{2}|\tau\frac{1}{2\pi}\lambda d\theta \tag{3}$$

$$= \frac{2r_0V\tau\lambda}{\pi}\int_0^{2\pi}|\sin\frac{\theta}{2}|d\theta = \frac{8r_0V\tau\lambda}{\pi} \tag{4}$$

Therefore, the expected number of nodes that have been inside the radio range of node S during its active period is $R_0 \leq \frac{8r_0V\tau\lambda}{\pi} + \pi r_0^2\lambda$. ■

B. Percolation probability

In this subsection, we study the percolation probability of an infrastructure-based MANET in the limit of large network size, where *in the limit of large network size* means that we increase the network area to infinity (i.e. let $L \to \infty$) while keeping other parameters, such as λ, V and r_0, constant. Note that the distance between INs also increases with L such that the ratio between the two values is kept constant as shown in Fig. 1. The percolation probability is defined in the following:

Definition 2. The *percolation probability* p_c of an infrastructure-based MANET is the probability that a piece of information broadcast from the INs is received by a *non-vanishingly-small* fraction of MNs in the steady state, in the limit of large network size.

We have the following result for the percolation probability.

Theorem 1. *Consider a network $\mathcal{G}(L, \lambda, M, V, \zeta, \tau)$, whose effective node degree is known to be R_0. The percolation probability is $p_c \leq 1 - (\frac{W(-R_0e^{-R_0})}{-R_0})^M$, where $W(.)$ is the Lambert W Function.*

Proof: We first consider the information dissemination process of a single IN, in which a node (say A) is a *child* of another node (say B) iff node A (which must be a susceptible node) receives the information from node B (which must be an infectious node). Denote by $\chi_N^1(k)$ the number of nodes at the k^{th} generation in the information dissemination process of a single IN. Then $\chi_N^1(k)$ can be modeled by a branching process whose expected number of children per node is smaller than R_0, because of the reduction of the fraction of susceptible nodes in the network as the information propagates. Next we introduce a Galton-Watson branching process [18], in which the expected number of children per node is R_0. Denoted by $\chi^1(k)$ the number of individuals at the k^{th} generation of the Galton-Watson branching process. As discussed earlier, we have $\chi_N^1(k) \leq \chi^1(k)$ in stochastic ordering[2] for all $k \geq 1$. Note that at the first generation, though there is no previous reduction of the fraction of susceptible nodes, the relation $\chi_N^1(1) \leq \chi^1(1)$ is still valid because an IN does not move over time so that the expected number of children of an IN can be smaller than R_0.

Define $q = \Pr(\lim_{k\to\infty}\chi^1(k) \to 0)$ to be the *extinction probability* of the Galton-Watson branching process $\chi^1(k)$. It can been shown (cf. [18, Theorem 6.5.1]) that $q = \frac{W(-R_0e^{-R_0})}{-R_0}$, where $W(.)$ is the Lambert W Function [19]. Further, define $q_N = \Pr(\lim_{k\to\infty}\chi_N^1(k) \to 0)$ to be the extinction probability of the branching process $\chi_N^1(k)$. Due to the stochastic ordering $\chi_N^1(k) \leq \chi^1(k)$, we have

$$q_N \geq q = \frac{W(-R_0e^{-R_0})}{-R_0} \tag{5}$$

Recall that we have M INs in the network. Denote by p_c' the probability that at least one of the M information dissemination processes, rooted at M INs respectively, does not become extinct. It can be shown that

$$p_c' \leq 1 - q_N^M \leq 1 - q^M \tag{6}$$

where the second inequality is due to Eq. 5 and the first inequality is due a spatial correlation between INs. The *spatial correlation between INs* arises if the Euclidean distance between two INs is small, for then two information dissemination processes may share a common set of susceptible nodes. As the information propagates, the reduction of the fraction of susceptible nodes caused by the information dissemination process of one IN can reduce the expected number of children per node in the information dissemination process of another IN. Therefore we have $p_c' \leq 1 - q_N^M$.

Further, the event of having an unbounded number of informed MNs in the limit of large network size (whose probability is p_c') is a necessary condition for having a non-vanishingly-small fraction of MNs in the limit of large network size. Therefore we have $p_c \leq p_c'$. ■

Remark 1. To improve the percolation probability of a given network $\mathcal{G}(L, \lambda, M, V, \zeta, \tau)$, one can reduce the spatial correlation between INs by reducing the number of susceptible nodes shared by the INs. Therefore we consider placing the INs in a way that the associated Voronoi cells have equal size, as introduced in Section III.

Remark 2. The above results suggest that an increase in the number of INs provides little improvement on the percolation probability when the density of MNs is either very low (such that $q = 1$) or very high (such that $q = 0$). This result coincides with the result for a static network [7]. However, our results further suggest that when the density of MNs is intermediate, an increase in the number of INs can improve the percolation probability significantly, which is further verified in Section V.

In addition, it can be shown that $1 - (\frac{W(-R_0e^{-R_0})}{-R_0})^M$ is 0 for $0 < R_0 \leq 1$ and it is monotonically increasing with

[2]Using stochastic ordering, we say $\chi_N^1(k) \leq \chi^1(k)$ iff $\Pr(\chi_N^1(k) > a) \leq \Pr(\chi^1(k) > a)$ for any number a.

respect to R_0 for $R_0 > 1$. Therefore the following corollary can be readily obtained using Lemma 1 and Theorem 1.

Corollary 2. *Consider a network $\mathcal{G}(L, \lambda, M, V, \zeta, \tau)$. The percolation probability is*

$$p_c \leq 1 - \left(\frac{W(-(\frac{8r_0V\lambda\tau}{\pi} + \pi r_0^2\lambda)e^{-\frac{8r_0V\lambda\tau}{\pi} - \pi r_0^2\lambda})}{-\frac{8r_0V\lambda\tau}{\pi} - \pi r_0^2\lambda}\right)^M$$

where $W(.)$ is the Lambert W Function.

The percolation probability provides the probability that a piece of information can be spread out to a significant fraction of MNs. In the next subsection, we quantify how many MNs can receive the information.

C. Expected fraction of informed MNs

Recall that a piece of information broadcast from INs is said to percolate if the information is received by a non-vanishingly-small fraction of MNs in the steady state.

Define z_p to be the expected fraction of informed MNs in the steady state given that the information percolates. Because $M \ll N$, the following result can be readily obtained based on Lemma 3 in [9].

Theorem 2. *Consider a network $\mathcal{G}(L, \lambda, M, V, \zeta, \tau)$, whose effective node degree is known to be R_0. If the information percolates, then the expected fraction of informed MNs in the steady state satisfies $z_p \leq 1 + \frac{1}{R_0}W(-R_0e^{-R_0})$, where $W(.)$ is the Lambert W Function.*

It is further shown in Section V that increasing the number of INs has little impact on z_p. On the other hand, increasing the number of INs can reduce the delay of the information dissemination process, which is studied in the next subsection.

D. Information dissemination delay

Consider that a piece of information is broadcast from the INs at time $t = 0$ using the SIR epidemic routing algorithm. Let $T(z)$ be the expected time when the fraction of informed MNs reaches z, for $0 < z < 1$. In this subsection, we provide a lower bound on the delay $T(z)$ by considering a SI in lieu of SIR epidemic routing algorithm.

It is worth noting that in a multi-hop network, a piece of information can be forwarded several hops in a short time interval if a journey exists in such time interval. The multi-hop forwarding is usually much faster than the store-carry-forward method because the latter method relies on the physical movement of MNs. Therefore the time delay of the information dissemination process allowing multi-hop forwarding is different and more realistic than that in the previous research of epidemics or MANETs based on a particular inter-meeting time or rate of transmission [11], [12], [16], as discussed in Section II. We take the multi-hop forwarding into consideration and provide the following result.

Theorem 3. *Consider a network $\mathcal{G}(L, \lambda, M, V, \zeta, \tau)$. A piece of information is broadcast from the INs at time $t = 0$ using the SIR epidemic routing algorithm. Let $T(z)$ be the expected time when the fraction of informed MNs reaches z, for $0 < z < 1$. Then $T(z) \geq \frac{\tau}{R_0+Nz_s} \ln \frac{(zN/M)-z}{1-z}$, where $z_s = (1 + \frac{1}{\lambda\pi r_0^2}W(-\lambda\pi r_0^2e^{-\lambda\pi r_0^2}))^2$ and $N = \lambda L^2$.*

Proof: Firstly we consider the multi-hop forwarding. Let z_m be the expected fraction of MNs that receive the information broadcast from an arbitrary node at an arbitrary time instant. By letting $V = 0$, Lemma 1 gives $R_0 \leq \lambda\pi r_0^2$. Then using Corollary 2 and Theorem 2, it can be shown that $z_m \leq (1 + \frac{1}{\lambda\pi r_0^2}W(-\lambda\pi r_0^2e^{-\lambda\pi r_0^2}))^2 \triangleq z_s$.

It is straightforward that the recovery mechanism reduces the number of infectious nodes hence slows down the information dissemination process. Therefore we can obtain a lower bound on $T(z)$ by ignoring the recovery mechanism: viz. considering a SI epidemic spreading where the expected number of nodes that come into the radio range of an arbitrary infectious node in a unit time interval is $\alpha = \frac{R_0+Nz_s}{\tau}$. Let $I(t)$ be the number of infectious nodes at time t, then it can be shown using the same method as that in Lemma 3 of [9]:

$$I'(t) = \alpha(1 - \frac{I(t)}{N})I(t) \tag{7}$$

Solving the above ODE, we have:

$$I(t) = \frac{\alpha \exp(C_1 + \alpha t)}{\frac{\alpha}{N}\exp(C_1 + \alpha t) + 1} \tag{8}$$

Because $I(0) = M$, we have $\exp(C_1) = \frac{1}{\alpha/M - \alpha/N}$. Then:

$$I(t) = \frac{\alpha \exp(\alpha t)}{\frac{\alpha}{N}\exp(\alpha t) + (\alpha/M - \alpha/N)} \tag{9}$$

$$= \frac{N}{1 + (N/M - 1)\exp(-\frac{R_0+Nz_s}{\tau}t)} \tag{10}$$

Letting $z = \frac{I(t)}{N}$, the above equation can be re-written as:

$$t(z) = \frac{\tau}{R_0 + Nz_s} \ln \frac{\frac{N}{M}z - z}{1 - z} \tag{11}$$

As discussed earlier, we have $T(z) \geq t(z)$. ■

Note that Theorem 3 can serve as a lower bound for other information dissemination methods for infrastructure-based MANETs, because the SI epidemic routing algorithm, like flooding, has been shown to have the lowest delay for information dissemination in MANETs [4], [10], [11].

V. Simulation results

In this section, we report on simulations to verify the accuracy of the analytical results. The simulations are conducted using a MANET simulator written in C++. MNs are deployed on a torus $(0, 800]^2$ following a homogeneous Poisson process with intensity $\lambda = 0.002$ users/m^2, which is the population density in Sydney city [20]. Speed V is 1.5m/s (typical human walking speed [21]) or 10m/s (typical vehicle moving speed [22]). The radio range r_0 is varied from 1 to 20. (The radio range of a Bluetooth device is around 10m [23].) The active period τ is set to be 100 or 200, where results using other values of τ show a similar trend. The INs are placed deterministically as shown in Fig. 1 and it remains our future

work to find the optimal placement of INs for the information dissemination in MANETs. Every point shown in the figures is the average value from 500 simulations, where the confidence interval is too small to be distinguishable and so is omitted.

In Fig. 3, the simulation result for percolation probability shows the probability that at least 5% of MNs receive the information broadcast from INs. Firstly, it can be seen in Fig. 3 that the analytical result well captures the impact of various parameters, e.g. nodal speed, radio range and the number of INs on the information dissemination process. Further, it can be seen in Fig. 3 (a) that a significant higher percolation probability can be achieved by a small number of additional INs (e.g. from 1 IN to 4 INs). On the other hand, it can be seen in Fig. 3 (b) that the introduction of more INs cannot improve the expected fraction of informed nodes in the steady state if the information percolates. This observation confirms our assertion given in Section IV-C. Further study reveals that if a MN (or a group of MNs) does not receive the information given that the information percolates elsewhere in the network, the MN (or the group of MNs) is usually geometrically separated from the informed nodes. However, a small number of additional INs has a small probability of bridging the geometric gap.

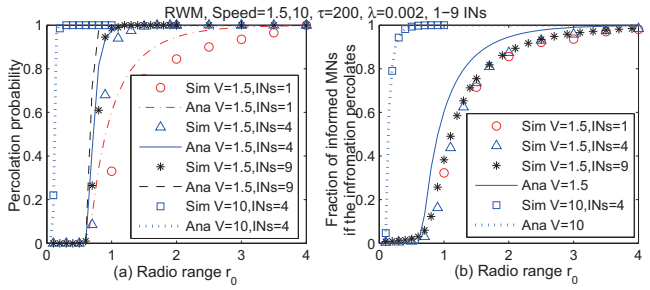


Fig. 3. Analytical (Ana) and simulation (Sim) results for (a) Percolation probability; (b) Expected fraction of informed MNs if the information percolates.

Fig. 4 shows the time delay for a piece of information to be received by 50% of MNs. It can be seen that our analytical result captures the impact on the delay of network parameters such as radio range, nodal speed or the number of INs.

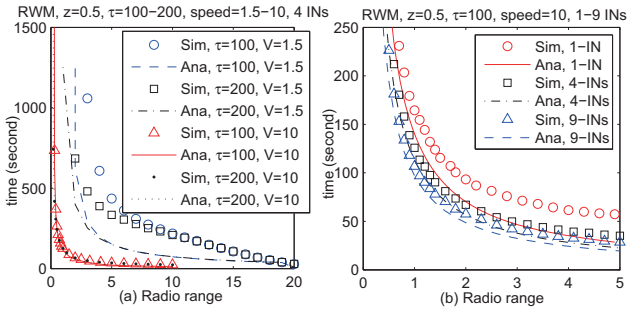


Fig. 4. Analytical (Ana) and simulation (Sim) results for the time delay for a piece of information to be received by 50% of MNs. The curves for V=10 overlap in (a), but the accuracy of the results for V=10 can be seen in (b).

VI. Conclusions and Future Work

In this paper, we study the dissemination of a piece of information broadcast from INs in an infrastructure-based MANET using a SIR epidemic routing algorithm. Analytical results are derived for the percolation probability, the expected fraction of informed MNs and the time delay. The accuracy of the analytical results is verified using simulations. In the future, we intend to study an infrastructure-based MANET driven by a real world trace.

References

[1] E. Daly and M. Haahr, "Social network analysis for routing in disconnected delay-tolerant MANETs," in *Proceedings of MobiHoc*, 2007, pp. 32–40.

[2] A. Vahdat and D. Becker, "Epidemic routing for partially-connected ad hoc networks," *Duke Tech Report CS-2000-06*, 2000.

[3] G. Mao and B. D. O. Anderson, "Graph theoretic models and tools for the analysis of dynamic wireless multihop networks," in *IEEE WCNC*, 2009, pp. 2828–2833.

[4] R. J. D'Souza and J. Jose, "Routing approaches in delay tolerant networks: A survey," *International Journal of Computer Applications*, vol. 1, no. 17, pp. 8–14, 2010.

[5] S. A. Bermudez and S. B. Wicker, "Partial connectivity of multi-hop two-dimensional finite hybrid wireless networks," in *Proc. IEEE International Conference on Communications, ICC*, 2010.

[6] Z. Zhang, S. C. Ng, G. Mao, and B. D. O. Anderson, "On the k-hop partial connectivity in finite wireless multi-hop networks," in *IEEE ICC*, 2011.

[7] O. Dousse, P. Thiran, and M. Hasler, "Connectivity in ad-hoc and hybrid networks," in *Proceedings. IEEE INFOCOM*, vol. 2, 2002, pp. 1079–1088.

[8] G. Zyba, G. M. Voelker, S. Ioannidis, and C. Diot, "Dissemination in opportunistic mobile ad-hoc networks: The power of the crowd," in *Proceedings IEEE INFOCOM*, 2011, pp. 1179–1187.

[9] Z. Zhang, G. Mao, and B. D. O. Anderson, "On the information propagation in mobile ad-hoc networks using epidemic routing," *accepted by Globecom2011*, pp. 1–6, 2011.

[10] R. Groenevelt, P. Nain, and G. Koole, "The message delay in mobile ad hoc networks," *Performance Evaluation*, vol. 62, no. 1-4, pp. 210 – 228, 2005.

[11] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, "Performance modeling of epidemic routing," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 51, no. 10, pp. 2867–2891, 2007.

[12] S. Zhou, L. Ying, and S. Tirthapura, "Delay, cost and infrastructure tradeoff of epidemic routing in mobile sensor networks," in *Proceedings of the 6th International Wireless Communications and Mobile Computing Conference*, 2010, pp. 1242–1246.

[13] M. Franceschetti and R. Meester, *Random networks for communication: From Statistical Physics to Information Systems*. Cambridge University Press, 2007.

[14] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communications and Mobile Computing*, vol. 2, no. 5, pp. 483 – 502, 2002.

[15] P. Nain, D. Towsley, B. Liu, and Z. Liu, "Properties of random direction models," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 3, 2005, pp. 1897– 1907.

[16] T. Britton, "Stochastic epidemic models: A survey," *Mathematical Biosciences*, vol. 225, no. 1, pp. 24–35, 2010.

[17] Z. Zhang, "Mobile ad-hoc networks," 2011. [Online]. Available: http://zijie.net/manet/

[18] P. Jagers, *Branching processes with biological applications*. Wiley, 1975.

[19] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W Function," *Advances in Computational Mathematics*, vol. 5, pp. 329–359, 1996.

[20] Demographia, "World urban areas: 7th annual edition," 2011. [Online]. Available: www.demographia.com/db-worldua.pdf

[21] N. Carey, "Establishing pedestrian walking speeds," *Karen Aspelin, Portland State University*, 2005.

[22] M. Rudack, M. Meincke, and M. Lott, "On the dynamics of ad hoc networks for inter vehicle communications (IVC)," in *Proc. ICWN*, 2002.

[23] A.-K. Pietilinen and C. Diot, "Experimenting with opportunistic networking," in *Proc. of the ACM MobiArch Workshop*, 2009.

On the Hop Count Statistics in Wireless Multi-hop Networks Subject to Fading

Zijie Zhang, *Student Member, IEEE,* Guoqiang Mao, *Senior Member, IEEE,*
and Brian D. O. Anderson, *Life Fellow, IEEE*

Abstract—Consider a wireless multi-hop network where nodes are randomly distributed in a given area following a homogeneous Poisson process. The hop count statistics, viz the probabilities related to the number of hops between two nodes, are important for performance analysis of the multi-hop networks. In this paper, we provide analytical results on the probability that two nodes separated by a known Euclidean distance are k hops apart in networks subject to both shadowing and small-scale fading. Some interesting results are derived which have generic significance. For example, it is shown that the locations of nodes three or more hops away provide little information in determining the relationship of a node with other nodes in the network. This observation is useful for the design of distributed routing, localization and network security algorithms. As an illustration of the application of our results, we derive the effective energy consumption per successfully transmitted packet in end-to-end packet transmissions. We show that there exists an optimum transmission range which minimizes the effective energy consumption. The results provide useful guidelines on the design of a randomly deployed network in a more realistic radio environment.

Index Terms—hop count, log-normal shadowing, Nakagami-m fading, energy consumption, wireless multi-hop networks.

1 Introduction

A wireless multi-hop network consists of a group of nodes that communicate with each other over wireless channels. The nodes in such a network operate in a decentralized and self-organized manner and each node can act as a relay to forward information toward the destination. Wireless multi-hop networks have large potential in military and civilian applications [1], [2].

There are three related probabilities characterizing the connectivity properties of such a multi-hop network. These are the probability that an arbitrary node is k hops apart from another arbitrary node, denoted by $\mathrm{r}(k)$; the probability that a node at an Euclidean distance x apart from another node is connected to that node in exactly k hops, denoted by $\mathrm{r}(k|x)$; and the spatial distribution of the nodes k hops apart from a designated node, denoted by $\mathrm{r}(x|k)$. These three probabilities are related through Bayes' formula and if one is computable, the other two will be computable using similar techniques. Therefore we call these problems collectively the *hop count statistics* problems. **We refer readers to Section 9 in the supplementary material for an extensive discussion on the use of the three probabilities in various applications.**

A major technical obstacle in the analysis of hop counts statistics is the so-called *spatial dependence problem* [3], [4]. The spatial dependence problem arises because in a wireless multi-hop network the event that a randomly chosen node is k hops apart from a particular node is not independent of the event that another randomly chosen node is i hops apart from the same node for $i \leq k$. It follows that an accurate analysis on the conditional probability $\mathrm{r}(k|x)$ needs to consider all previous hops, which makes the analysis complicated. This technical hurdle caused by the spatial dependence problem is recognized in the literature but remains unsolved [4]. In this paper, a significant improvement on the accuracy of computing $\mathrm{r}(k|x)$ is achieved by considering the positions of previous two hop nodes, compared to the results considering the positions of previous one hop nodes only. Further, we show that considering the positions of previous two hops nodes is enough to provide an accurate estimate of $\mathrm{r}(k|x)$. A detailed explanation of the spatial dependence problem is given in Section 3.1.

Further, much previous research (e.g. [5]–[8]) establishes results based on the assumption that the network is *connected*, i.e. there is at least one path between every pair of nodes. However in many wireless multi-hop network applications, it is not only impractical (due to the randomness of node deployment, a complex radio environment, or the high node density required for a large scale network to be connected [9]) but also unnecessary to require every node to be connected to every other node. For example, in applications which are not life-critical, e.g. habitat monitoring or environmental mon-

- *Z. Zhang and G. Mao are with the School of Electrical and Information Engineering, University of Sydney. G. Mao is also with National ICT Australia, Sydney. E-mail: {zijie.zhang, guoqiang.mao}@sydney.edu.au.*
- *B. D. O. Anderson is with the Research School of Information Sciences and Engineering, Australian National University and National ICT Australia, Canberra. E-mail: brian.anderson@anu.edu.au.*

This work is partially supported by Australian Research Council (ARC) Discovery project DP110100538, and by the Air Force Research Laboratory, under agreement number FA2386-10-1-4102. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

itoring, having a few disconnected source-destination pairs will not cause statistically significant change in the monitored parameters [10]. In addition, there is a downside on the capacity to have a high node density required for a connected network. Gupta and Kumar [11] showed that in a connected network, as the number of nodes per unit area n increases, the throughput per source-destination pair decreases approximately as $1/\sqrt{n}$. It was further pointed out in [12] that significant energy savings can be achieved by requiring most nodes but not all nodes to be connected. Therefore we consider the more realistic and practical scenario that the network is not necessarily connected. Previous results established on the basis of a connected network actually form special cases of the problem examined in this paper.

The main contributions of this paper are:

- Firstly, in a network with nodes distributed in a finite area following a homogeneous Poisson process, we derive the probability that two nodes separated by a known Euclidean distance x are k hops apart, i.e. $\mathrm{r}(k|x)$, considering both shadowing and small-scale fading and using a distributed routing algorithm, i.e. greedy forwarding;
- Secondly, we analyze the impact of the spatial dependence problem on the accuracy of $\mathrm{r}(k|x)$;
- Thirdly, considering a sparse network in which there is not necessarily a path between any pair of nodes, we derive the probability distribution of the number of hops traversed by packets before being dropped if the transmission is unsuccessful. (An end-to-end transmission is *unsuccessful* if the packet sent from a source to a destination has to be dropped at an intermediate node because it is unable to find a next-hop node.)
- As an application of the results, we derive the effective energy consumption per successfully transmitted packet in end-to-end packet transmissions. We show that there exists an optimum transmission range which minimizes the effective energy consumption. Further, the impacts of unreliable link, node density and path loss exponent on the energy consumption is included in the analysis.

The results in this paper provide a more complete understanding on the properties of wireless multi-hop networks in a more realistic and practical setting.

The rest of this paper is organized as follows: Section 8, in the supplementary material, reviews the related work. Section 2 introduces the network models and some definitions. The analysis of the hop count statistics and the end-to-end energy consumption under the unit disk communication model is given in Section 3, followed by the analysis on the impact of the spatial dependence problem on the accuracy of the hop count statistics in Section 4. In Section 5, we further include the log-normal shadowing and small-scale fading in the analysis. The simulation results and discussions are given in Section 6. Finally Section 7 concludes this paper and proposes possible future work.

2 System Model

2.1 Network model

In this paper, we consider a wireless multi-hop network where nodes are identically and independently distributed (i.i.d.) in a square according to a homogeneous Poisson point process with a known intensity ρ.

We consider that every node has the same transmission power. The simplest radio propagation model is the unit disk communication model. Under the unit disk model, the power attenuates with the Euclidean distance x from a transmitter like $x^{-\eta}$, where η is the path loss exponent. The path loss exponent can vary from 2 in free space to 6 in urban areas [18]. The received signal strength (RSS) at a receiver separated by Euclidean distance x from the transmitter is $P_u(x) = CP_t x^{-\eta}$, where C is a constant, P_t is the transmission power. A transmission is successful iff the RSS exceeds a given threshold $P_{\min}$. Therefore the required transmission power P_t allowing a transmission range r_0 is $P_t = C_1 r_0^{\eta}$, where $C_1 = P_{\min}/C$.

The unit disk model is simple but unrealistic. In reality the RSS may have significant variations around the mean value, because of both large scale variation (i.e. shadowing) and small-scale fading. Considering a typical type of shadowing, i.e. the *log-normal shadowing model* [18] the RSS attenuation (in dB) follows a normal distribution with respect to the distance x between transmitter and receiver: $10 \log_{10}(P_l(x)/CP_t x^{-\eta}) \sim Z$, where $P_l(x)$ is the RSS in the log-normal shadowing model and Z is a zero-mean Gaussian distributed random variable with standard deviation σ. When $\sigma = 0$ the model reduces to the unit disk model. In practice the value of σ is often computed from measured data and can be as large as 12 [18]. Denote by $q(z)$ the pdf (probability density function) of the shadowing fades; then: $q(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{z^2}{2\sigma^2})$. As widely used in the literature [18], [20], [25], we assume that the shadowing fades Z between all pairs of transmitting node and receiving node are i.i.d. and the link is symmetric. **The limitation of the assumption of the independence between links is discussed in Section 10 in the supplementary material.**

Shadowing makes the RSS vary around its mean value over space, while the small-scale fading makes the RSS vary around its mean value over time. In this paper, we consider a generic model of small-scale fading, i.e. the *Nakagami-m fading* [26]. By choosing different values for the parameter m in the Nakagami-m fading model, the results easily include several widely used fading distributions as special cases, e.g. Rayleigh distribution (by setting $m = 1$) and one-sided Gaussian distribution (by setting $m = 1/2$) [26]. Subject to Nakagami-m fading, the RSS per symbol, ω, is distributed according to a Gamma distribution given by the following pdf [26]:

$$\zeta(\omega) = \frac{m^m \omega^{m-1}}{\bar{\omega}\Gamma(m)} \exp(-\frac{m\omega}{\bar{\omega}}), \quad \omega \geq 0 \tag{1}$$

where $\Gamma(.)$ is the standard Gamma function and $\bar{\omega} = P_l(x)$ is the mean RSS (over time), which is determined by path loss and shadowing.

We firstly conduct the analysis in the unit disk model, then we introduce the analysis in the more realistic *Log-normal-Nakagami* model, which takes into account statistical variations of RSS around the mean value due to both log-normal shadowing and Nakagami-m fading.

2.2 Per hop energy consumption

Assume that the time spent on transmitting a packet of unit size over a single hop is a constant T_t, and all nodes transmit at the same power P_t which results in a transmission range (without shadowing) of r_0 in the unit disk model. Therefore the energy consumed in transmitting a packet over a single hop is:

$$Eng(r_0) = \frac{T_t P_t + Eng_c}{1 - \alpha(r_0)} = \frac{C_2 r_0^{\eta} + Eng_c}{1 - \alpha(r_0)} \qquad (2)$$

where $C_2 = C_1 T_t$ is a constant, Eng_c is another constant which includes the processing power consumption and receiving power consumption in each node and $\alpha(r_0) = \frac{2WN_b}{W+1^{\ 2}+2WN_b}$ is the packet error rate [27], W_{min} is the minimum contention window size and N_b is the average node degree: $N_b = \rho\pi r_0^2$. Packet collision can increase the energy consumption, due to the consequent re-transmission of a packet, especially when the transmission range is large. To illustrate this effect, we implemented simulations (in Section 6) using the parameters shown in [27], i.e. $W_{min} = 64$. The values of C_2 and Eng_c are dependent on hardware specifications, where some typical values can be found in [28].

2.3 Routing algorithm

In addition to the impact of fading on the wireless channel between two nodes, fading also affects the performance of higher layer protocols. The impact of fading on higher layer protocols remains to be fully investigated [22]. In this paper, we consider the cross-layer issues by analyzing the performance of a wireless multi-hop network using the greedy forwarding routing algorithm, as a typical example of distributed routing algorithms. The GF routing belongs to the category of geographic routing algorithms and is a widely used routing algorithm for wireless multi-hop networks. Using GF, each node makes routing decisions independently of other nodes by using its own location information, the location information of its neighboring nodes and the locations of the source and the destination. GF has shown great potential in wireless multi-hop networks because of its distributed nature, low control overhead and capability of adapting to dynamic network topologies [14]. The area has attracted significant research interest, e.g. [5]–[8].

We consider a basic GF algorithm that operates following two rules [7]: 1) Every node tries to forward the packet to the node within its transmission range which is closest to the destination. 2) A packet will be dropped if a node cannot find a next-hop neighbor that is closer to the destination than itself, and hence the transmission becomes unsuccessful. **Moreover in the case of ties, viz. more than one node have the same Euclidean distance to the destination, an arbitrary one of those nodes can be chosen as the next hop node without affecting the results of our analysis. This is because the way to settle ties does not affect the probability distribution of the remaining distance to the destination at each hop, which is the quantity used to derive our results as shown in Section 3.2.**

Note that a number of complicated recovery algorithms have been proposed to route a packet around the routing void [29]. The quality of the path established by a greedy forwarding algorithm can be measured by the stretch factor, which quantifies the difference between a particular path and the shortest path [29]. By studying the stretch factor, it is shown in [29] that a basic GF algorithm can successfully find short routing paths in sensing-covered networks, without complex recovery algorithms. For analytical tractability and generality of the results, we consider the basic greedy forwarding algorithm without any recovery algorithm, as in [5]–[8].

2.4 Definitions of some terms

Denote by $E_s[k_s|x_0]$ the expected number of hops for a packet to reach the destination, conditioned on the Euclidean distance between source and destination being x_0 *and* the transmission being successful. For convenience, throughout this paper we use *conditioned on* x_0 for *conditioned on the Euclidean distance between source and destination being* x_0. Denote by $E_u[k_u|x_0]$ the expected number of hops traversed by a packet before it is dropped due to the nonexistence of a next hop node closer to the destination, conditioned on x_0 *and* the transmission being unsuccessful (in this case x_0 is the distance between the source and the *intended* destination).

It is worth noting that with the assumption that the network is connected, as used in say [7], the number of hops between two nodes increases as the transmission range (hence average node degree) decreases. In terms of energy consumption, the assumption of a connected network results in a misleading conclusion that a smaller transmission range is always better. This conclusion is misleading because the probability of having a multi-hop path between two nodes decreases as the transmission range decreases and this important fact was not considered. Consequently, this conclusion is in sharp contrast with the result obtained in this paper considering the possibility of disconnected networks that there exists an optimum transmission range that minimizes the energy consumption, as shown in Fig. 7. Our analysis does not rely on the assumption that the network is connected.

Let's consider a network with a total of N distinct source and destination pairs, where each source is separated from the associated destination by Euclidean distance x_0. Each source transmits a packet of unit size

to the associated destination. Therefore there are a total of N packets transmitted. Assume M ($M \leq N$) packets can reach their respective destinations successfully.

Define $Eng_{eff}(r_0|x_0)$ to be the effective energy consumption per successfully transmitted packet for any pair of nodes separated by Euclidean distance x_0, viz $Eng_{eff}(r_0|x_0)$ is the total energy spent on transmitting all packets divided by the number of successfully received packets:

$$\begin{aligned} & Eng_{eff}(r_0|x_0) \qquad (3) \\ = & \frac{MEng(r_0)E_s[k_s|x_0] + (N-M)Eng(r_0)E_u[k_u|x_0]}{M} \\ = & Eng(r_0)\frac{\phi_s(x_0)E_s[k_s|x_0] + (1-\phi_s(x_0))E_u[k_u|x_0]}{\phi_s(x_0)} \end{aligned}$$

where $\phi_s(x_0) := M/N$ is the probability of successful transmission between any pair of nodes separated by x_0.

In a network where the transmission range without shadowing and fading is r_0, given the distribution of the Euclidean distance between any pair of nodes $f(x_0)$, examples of which are given in [30], the average effective energy consumption is:

$$Eng_{eff}(r_0) = \int Eng_{eff}(r_0|x_0)f(x_0)dx_0 \qquad (4)$$

The effective energy consumption is a measure of the energy spent on each successfully transmitted packet. A lower Eng_{eff} means a higher energy efficiency. We use Eng_{eff} as the metric to investigate the energy efficiency in end-to-end packet transmissions.

3 Analysis in the Unit Disk Model

In this section, we analyze the hop count statistics (in particular the probability $\mathrm{r}(k|x)$) and the effective energy consumption under the unit disk model. We start with the calculations of the probability that two arbitrary nodes are k hops apart for $k \geq 3$ using GF. The analysis for $k = 1, 2$ is straightforward.

3.1 Spatial dependence problem

Before going into the analysis, we introduce the spatial dependence problem in the analysis of hop count statistics using the unit disk model as an example. The same problem also exists in other models. Generally, there are two types of spatial dependence problems.

First, it can be shown that the event that a randomly chosen node is a k^{th} hop node (S_k) from a randomly chosen source node (S) is not independent of the event that another randomly chosen node is a i^{th} hop node for $1 \leq i < k$. Denote by $C(S_k, r)$ the disk centered at S_k with radius r. As shown in the example in Fig. 1(a), the fact that S_{k-1} is a $k-1^{th}$ hop node from a source node S (not shown in the figure) implies that there is at least one node in the area $C(S_{k-3}, r_0) \cap C(S_{k-1}, r_0)$. On the other hand S_k is a k^{th} hop node from S implies that there is no node in the area $C(S_{k-3}, r_0) \cap C(S_k, r_0)$,

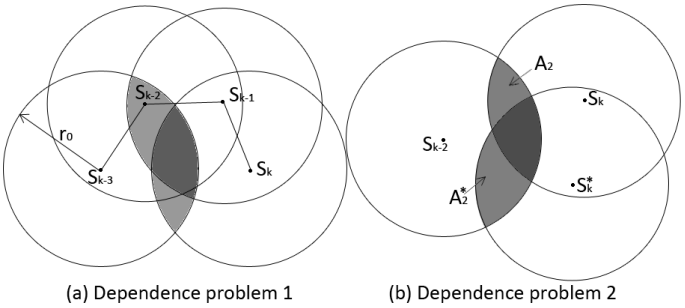


Fig. 1: Illustration of the spatial dependence problems in the hop count statistics using a unit disk model. S_k is the k^{th} hop node, where r_0 is the transmission range.

otherwise S_k will become a $k-2^{th}$ hop node. These two areas overlap which means that the event that S_k is a k^{th} hop node and the event that S_{k-1} is a $k-1^{th}$ hop node are not independent.

Second, it can be shown that the event that a randomly chosen node is a k^{th} hop node from S is not independent of the event that another randomly chosen node is a k^{th} hop node from S. As shown in the example in Fig. 1(b), S_k is a k^{th} hop node from S implies that there is at least one node (the $k-1^{th}$ hop node) in the area $A_2 = C(S_{k-2}, r_0) \cap C(S_k, r_0)$. Another node S_k^* is a k^{th} hop node from the same S implies that there is at least one node (the $k-1^{th}$ hop node) in the area $A_2^* = C(S_{k-2}, r_0) \cap C(S_k^*, r_0)$. These two areas overlap which means that the event that S_k is a k^{th} hop node and the event that S_k^* is a k^{th} hop node are not independent. In this paper, a significant improvement on the accuracy of $\mathrm{r}(k|x)$ is shown by reducing the inaccuracy associated with the first type of spatial dependence problem. **The second type of spatial dependence problem can be handled by a similar technique used in this paper. Specifically, when considering the area covered by the transmission range of a k^{th} hop node, we need to consider the overlapping of the area covered by the transmission range of the k^{th} hop node and the area covered by the transmission range of other k^{th} hop nodes. However, it can be seen in Section 7 that the result is fairly accurate after a proper handling of the first type of spatial dependence problem, so that specific handling of this second spatial dependence problem is effectively not warranted.**

3.2 Distribution of the remaining distance

Denote by $A(x, r_1, r_2)$ the intersectional area of two disks with distance x between centers and radii r_1 and r_2 respectively. The size of the area is [31]:

$$A(x, r_1, r_2) = \qquad (5)$$
$$\begin{cases} \mathrm{m} \quad (\pi r_1^2, \pi r_2^2), \quad \text{for } x \leq |r_1 - r_2| \\ r_1^2 \arccos(\frac{x^2+r_1^2-r_2^2}{2xr_1}) + r_2^2 \arccos(\frac{x^2+r_2^2-r_1^2}{2xr_2}) \\ -\frac{1}{2}\sqrt{[(r_1+r_2)^2 - x^2][x^2 - (r_1-r_2)^2]}, \\ \quad \text{for } |r_1 - r_2| < x < r_1 + r_2 \\ 0, \quad \text{otherw se} \end{cases}$$

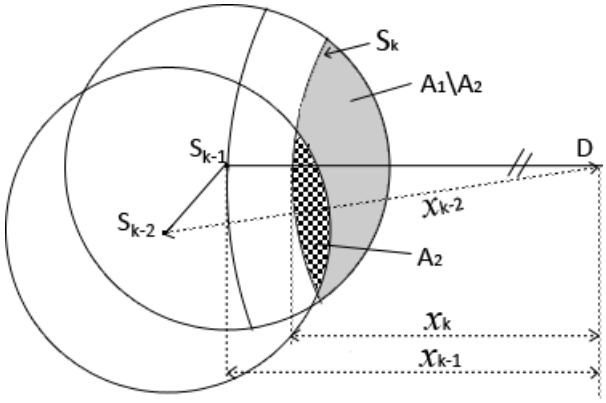


Fig. 2: Possible positions for the node at the k^{th} hop, denoted by S_k, are located on the arc, considering the positions of S_{k-1} and S_{k-2}, which are the nodes at the $k-1^{th}$ hop and $k-2^{th}$ hop from the source respectively. A_1, A_2, x_k, x_{k-1} and x_{k-2} are described in the following text.

Define x_k to be the remaining Euclidean distance between the k^{th} hop node (S_k) and the destination (D). Define $A_1 = A(x_{k-1}, r_0, x_k)$ to be the intersectional area of the disks $C(S_{k-1}, r_0)$ and $C(D, x_k)$. Similarly we have $A_2 = A(x_{k-2}, r_0, x_k)$. Next we record the form of $\frac{\partial A\ x,r_1,r_2}{\partial r_2}$, which will be used later. For $|r_1 - r_2| < x < r_1 + r_2$:

$$\begin{aligned} \frac{\partial A(x,r_1,r_2)}{\partial r_2} &= \frac{-r_1^2}{\sqrt{1-S^2}}(\frac{\partial S}{\partial r_2}) + 2r_2 \arccos(T) \\ &\quad + r_2^2(\frac{-1}{\sqrt{1-T^2}})(\frac{\partial T}{\partial r_2}) - \frac{1}{4\sqrt{W}}(\frac{\partial W}{\partial r_2}) \\ \text{where} \quad S &= \frac{x^2+r_1^2-r_2^2}{2xr_1}, \quad T = \frac{x^2+r_2^2-r_1^2}{2xr_2}, \\ W &= ((r_1+r_2)^2 - x^2)(x^2-(r_1-r_2)^2) \end{aligned} \tag{6}$$

Define $f(x_k, k|x_0)$ to be the joint pdf of the remaining Euclidean distance to the destination from S_k being x_k *and* the packet having been successfully forwarded k hops, conditioned on x_0. Due to the spatial dependence problem, $f(x_k, k|x_0)$ depends on the remaining distances of all previous hop nodes, i.e. x_{k-1}, x_{k-2}, ..., x_0. We consider no more than two previous hops and the justification is given in Section 4.

Define $g(x_k, k|x_{k-1}, x_{k-2}, k-1)$ to be the joint pdf of the remaining Euclidean distance to the destination at S_k being x_k *and* the packet having been successfully forwarded k hops, conditioned on $\mathcal{B}$, where $\mathcal{B}$ is the event that the remaining distances at S_{k-1} and S_{k-2} are x_{k-1} and x_{k-2} respectively *and* the packet has been successfully forwarded $k-1$ hops. (Note that a packet has been successfully forwarded $k-1$ hops necessarily means that it has been successfully forwarded i hops for $i \leq k-1$.) Accordingly define the cdf (cumulative distribution function) of the remaining distance at the k^{th} hop node to be $\mathrm{r}(X_k \leq x_k, k|x_{k-1}, x_{k-2}, k-1)$. Ignoring the boundary effect, whose impact will be discussed in detail later, the cdf is equal to the probability that there is at least one node in area $A_1 \setminus A_2$ as indicated by the uniform-shaded area in Fig. 2. The area A_2 needs to be excluded because if there is a node in this area, that node will be closer to the destination than S_{k-1}, which violates the condition that S_{k-1} is the $k-1^{th}$ hop node using GF. We approximate the size of $A_1 \setminus A_2$ by $A_1 - A_2$. This approximation will greatly simplify the calculation while giving a sufficiently accurate result, as validated in Section 6. Due to space limitation, we omitted analytical studies on the accuracy of the approximation. Then:

$$\begin{aligned} &\mathrm{r}(X_k \leq x_k, k|x_{k-1}, x_{k-2}, k-1) \\ &= 1 - \mathrm{e}\ \mathrm{p}(-\rho(A(x_{k-1}, r_0, x_k) - A(x_{k-2}, r_0, x_k))) \end{aligned} \tag{7}$$

For any two nodes close to the border, the intersectional area of the transmission ranges of the two nodes may be partially located outside the network area, which causes an error in computing the size of the area $A_1 \setminus A_2$ in Eq. 7. This effect is due to the boundary effect. Ignoring the boundary effect may generally cause an overestimation on the size of $A_1 \setminus A_2$, hence an overestimation on the probability of finding the next hop node. However, simulation results in Section 6 show that the boundary effect has very limited impact on the accuracy of the analytical results.

Taking the derivative of the cdf with respect to x_k, we have:

$$\begin{aligned} &g(x_k, k|x_{k-1}, x_{k-2}, k-1) \\ &= \frac{\partial\ \ \mathrm{r}(X_k \leq x_k, k|x_{k-1}, x_{k-2}, k-1)}{\partial x_k} \\ &= \rho\left(\frac{\partial A(x_{k-1}, r_0, x_k)}{\partial x_k} - \frac{\partial A(x_{k-2}, r_0, x_k)}{\partial x_k}\right) \\ &\times \mathrm{e}\ \mathrm{p}(-\rho(A(x_{k-1}, r_0, x_k) - A(x_{k-2}, r_0, x_k))) \end{aligned} \tag{8}$$

where the partial differentiations are given by Eq. 6.

Define $h(x_k, x_{k-1}, k|x_0)$ to be the joint pdf of the remaining Euclidean distances at the k^{th} hop node and $k-1^{th}$ hop node being x_k and x_{k-1} respectively *and* the packet having been successfully forwarded k hops, conditioned on x_0.

For $k = 1$, it is straightforward that:

$$f(x_1, k=1|x_0) = \rho\frac{\partial A(x_0, r_0, x_1)}{\partial x_1} e^{-\rho A\ x_0, r_0, x_1} \tag{9}$$

For convenience, $f(x_1, k = i|x_0)$ is denoted by $f(x_1, i|x_0)$ hereafter. Based on the above result, for $k = 2$ we have:

$$h(x_2, x_1, 2|x_0) = g(x_2, 2|x_1, x_0, 1) f(x_1, 1|x_0) \tag{10}$$

For $k > 2$, $h(x_k, x_{k-1}, k|x_0)$ can be calculated recursively:

$$\begin{aligned} h(x_k, x_{k-1}, k|x_0) &= \int_{r_0}^{x_0} g(x_k, k|x_{k-1}, x_{k-2}, k-1) \\ &\quad h(x_{k-1}, x_{k-2}, k-1|x_0) dx_{k-2} \end{aligned} \tag{11}$$

Finally for $k > 1$ we have:

$$f(x_k, k|x_0) = \int_{r_0}^{x_0} h(x_k, x_{k-1}, k|x_0) dx_{k-1} \tag{12}$$

3.3 Hop count statistics

Define $\mathrm{r}(k|x_0)$ to be the probability that the destination can be reached at the k^{th} hop conditioned on x_0. The destination can be reached at the k^{th} hop if the $k-1^{th}$ hop node is within the transmission range of the destination. Therefore:

$$\mathrm{r}(k|x_0) = \int_0^{r_0} f(x_{k-1}, k-1|x_0)dx_{k-1} \quad (13)$$

3.4 Results for successful transmissions

Denote by $\mathrm{r}_s(k_s|x_0)$ the conditional probability that a packet can reach its destination at the k_s^{th} hop, conditioned on x_0 *and* the transmission being successful. It follows that $\mathrm{r}(k_s|x_0) = \mathrm{r}_s(k_s|x_0)\phi_s(x_0)$, and:

$$\begin{aligned}\sum_{k_s=1}^{\infty} k_s \ \mathrm{r}(k_s|x_0) &= \phi_s(x_0) \sum_{k_s=1}^{\infty} k_s \ \mathrm{r}_s(k_s|x_0) \\ &= \phi_s(x_0) E_s[k_s|x_0] \quad (14)\end{aligned}$$

where $\phi_s(x_0)$ and $E_s[k_s|x_0]$ are defined in Section III.D. In reality, an upper bound on k_s can be found beyond which $\mathrm{r}(k_s|x_0)$ is 0. So Eq. 14 and other similar equations only need to be computed for a finite range of k_s.

An end-to-end packet transmission is successful if a packet can reach the destination at any number of hops. Therefore:

$$\phi_s(x_0) = \sum_{k_s=1}^{\infty} \mathrm{r}(k_s|x_0) \quad (15)$$

3.5 Results for unsuccessful transmissions

Define $\phi_u(k|x_0)$ to be the probability of a packet having been successfully forwarded k hops from the source toward the destination x_0 apart, but *not* reaching the destination in k hops, which distinguishes $\phi_u(k|x_0)$ from $\mathrm{r}(k|x_0)$. Therefore:

$$\phi_u(k|x_0) = \int_0^{x_0} f(x_k, k|x_0)dx_k \quad (16)$$

Based on the example introduced in Section III.D, we further assume that only M_k out of N packets reach the k^{th} hop nodes. Then $\phi_u(k|x_0) = M_k/N$. At the next hop, there are three possibilities for each of these M_k packets: 1) a packet reaches the destination at the next hop; 2) a packet makes another hop without reaching the destination; 3) the packet is dropped. Let W_{k+1} and M_{k+1} be the number of packets for which the first and second possibilities apply.

Define $\psi(k_u|x_0)$ to be the probability of the packets being dropped at the k_u^{th} hop. Then:

$$\begin{aligned}\psi(k_u|x_0) &= \frac{M_{k_u} - M_{k_u+1} - W_{k_u+1}}{N} \\ &= \phi_u(k_u|x_0) - \phi_u(k_u+1|x_0) - \mathrm{r}(k_u+1|x_0) \quad (17)\end{aligned}$$

The average number of hops for unsuccessful transmissions between a source and a destination separated x_0 apart is the expected value of k_u whose pdf is given by $\psi(k_u|x_0)$. Similar to the way to derive Eq. 14, we have: $\sum_{k_u=1}^{\infty} k_u\psi(k_u|x_0) = (1-\phi_s(x_0))E_u[k_u|x_0]$.

Given the above analysis, the effective energy consumption can be computed using Eq. 4, which is shown in Section 6. Further, the above results can also be useful in the analysis of delay, throughput or reliability of end-to-end packet transmissions [8], [20], as well as localization [8], [32], which is left as our future work.

4 Impact of Spatial Dependence Problem

In this paper, we considered that the remaining distance at the k^{th} hop node (S_k) depends on the remaining distance at previous two hops nodes (S_{k-1} and S_{k-2}). Due to the spatial dependence problem, it can be shown that correct analysis of the hop count statistics requires all previous hops to be considered, but the calculation is more complicated than if an independence assumption is made. Previous research, e.g. [7], usually considered the dependence on only previous one hop. In this section we study the impact of the spatial dependence problem on the accuracy of the $\mathrm{r}(k|x)$.

Define A_m to be the intersectional area of the disk centered at S_{k-m} with radius r_0 and the disk centered at D with radius x_k. Therefore the precise area that should be considered in the calculation of Eq. 7 is $A = A_1 \setminus (A_2 \cup A_3 \cup ... \cup A_k)$ instead of $A_1 \setminus A_2$.

Consider only previous one hop, then: $A \approx A_1$. Consider only previous two hops, then: $A \approx A_1 \setminus A_2 = A_1 - A_1 \cap A_2$. Consider only previous three hops, then:

$$\begin{aligned}A &\approx A_1 \setminus (A_2 \cup A_3) = A_1 - A_1 \cap (A_2 \cup A_3) \\ &= A_1 - A_1 \cap A_2 - A_1 \cap A_3 + A_1 \cap A_2 \cap A_3 \quad (18)\end{aligned}$$

The underlined terms are the additional terms introduced when considering one more previous hop. In considering the previous m hops instead of previous $m-1$ hops, the improvement is bounded by a term determined by $A_1 \cap A_m$. Furthermore, it is evident that $x_{k-1} < x_{k-2} < ... < x_0$. Therefore $A_1 > A_2 > ... > A_k$ and the size of $A_1 \cap A_m$ is dominated by the size of A_m.

Define $h(x_k, x_{k-m}, k|x_0)$ to be the joint pdf of the remaining Euclidean distances at the k^{th} hop node and $k-m^{th}$ hop node being x_k and x_{k-m} respectively *and* the packet having been successfully forwarded k hops, conditioned on x_0. Then the expected size of A_m at the k^{th} hop can be calculated by:

$$\begin{aligned}E[A_m, k|x_0] &= \int_0^{x_0} \int_{x_k}^{x_k+r_0} A(x_{k-m}, r_0, x_k) \\ &\quad h(x_k, x_{k-m}, k|x_0)dx_{k-m}dx_k \quad (19)\end{aligned}$$

For $m=1$, $h(x_k, x_{k-1}, k|x_0)$ can be calculated using Eq. 11. For $m=2$, we have:

$$\begin{aligned}h(x_k, x_{k-2}, k|x_0) &= \int_{r_0}^{x_0} g(x_k, k|x_{k-1}, x_{k-2}, k-1) \\ &\quad h(x_{k-1}, x_{k-2}, k-1|x_0)dx_{k-1} \quad (20)\end{aligned}$$

For $m \geq 3$, the calculation becomes more complicated. But approximately $h(x_k, x_{k-m}, k|x_0) \approx$

$f(x_k, k|x_0)f(x_{k-m}, k-m|x_0)$, where $f(x_k, k|x_0)$ is given by Eq. 12. This approximation is valid because the distance between S_k and S_{k-m} generally increases as m increases, hence the size of the overlapping area decreases. Therefore the correlation between $f(x_k, k|x_0)$ and $f(x_{k-m}, k-m|x_0)$ reduces as m increases.

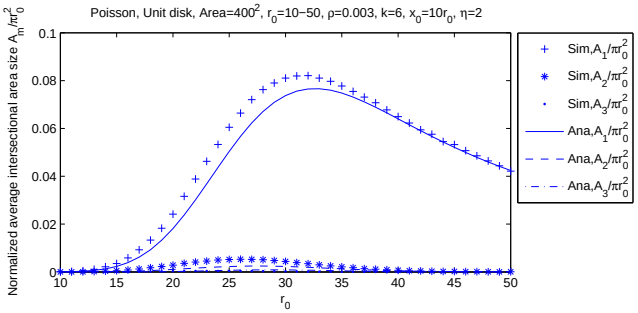


Fig. 3: Simulation (Sim) and analytical (Ana) results for the normalized average intersectional area size in the unit disk model. A_m is the intersectional area of the disk centered at S_{k-m} with radius r_0 and the disk centered at D with radius x_k.

Based on the approach introduced above, Fig. 3 shows the results for the average size of A_m when the source node and the destination node are separated by distance $x_0 = 10r_0$. The simulation parameters are introduced in Section 6. It is evident that the size of A_m, $m \geq 3$, are negligibly small (less than 1% of the size of the area covered by the transmission range) compared to the size of A_1 and A_2. It validates the claim that the improvement made by taking previous m hops into consideration will be marginal for $m \geq 3$, which explains our choice of considering two previous hops only.

Our results suggest that the accuracy of the analysis on $\mathrm{r}(k|x)$ can be significantly improved by considering previous two hops (compared to considering previous one hop only). However, moving beyond two hops results in marginal improvement in accuracy of the analysis. Therefore, the conclusion can be drawn that the locations of nodes three or more hops away provide little information for a node to determine its geometric relationship with other nodes. This conclusion provides analytical support for observations, to this point unsupported by analysis, in routing, localization and network security that taking into account the (location or link status) information of two-hops neighbors can significantly improve the routing [33] (respectively localization [34], network security [35]) performance compared with using one-hop neighborhood information only. However beyond two hops, taking into account more neighborhood information only has marginal impact. Therefore many distributed routing, localization and network security protocols use two-hop neighborhood information.

5 Analysis in the Log-normal-Nakagami Model

The technique to incorporate the impacts of both shadowing and small-scale fading is through the use of the random split property of a Poisson process.

5.1 Random split of a Poisson process

First, we introduce a random variable named the *Nakagami fades* Ω_0 which follows the Gamma distribution with mean 1. Therefore the pdf of Ω_0 is

$$\zeta_0(\omega_0) = \frac{m^m \omega_0^{m-1}}{\Gamma(m)} \mathrm{e}\ \mathrm{p}(-m\omega_0), \quad \omega_0 \geq 0 \tag{21}$$

where m is introduced in Section 2.1.

It can be shown that the random variable $P_l(x)\Omega_0$ follows the Gamma distribution with mean $P_l(x)$, where $P_l(x) = CP_t x^{-\eta} 10^{Z/10}$ is the RSS given by the log-normal shadowing model introduced in Section 2.1. Then in the Log-normal-Nakagami model, the RSS at a receiver at distance x from the transmitter is $P_N(x) = P_l(x)\Omega_0 = CP_t x^{-\eta} 10^{Z/10}\Omega_0$, where Z is a zero-mean Gaussian distributed random variable and Ω_0 is a Gamma distributed random variable with mean 1.

According to the random split property of a Poisson process [36], the subset of nodes whose RSS from a particular transmitting node with shadowing fades $Z \in [z, z+dz]$ *and* Nakagami fades $\Omega_0 \in [\omega_0, \omega_0 + d\omega_0]$ are i.i.d. following a Poisson process with intensity $\rho q(z) dz \zeta_0(\omega_0) d\omega_0$. Via the splitting of the Poisson process, we can study the sub-process by the same technique used in the unit disk model.

Remark 1: The aforementioned technique can be extended to other communication models (e.g. the class of random connection models [37]). In a random connection model [37], two arbitrary nodes separated by Euclidean distance x are directly connected with probability $\gamma(x)$, where $\gamma(x)$ satisfies two conditions: 1) the probability is a non-increasing function mapping from the positive real numbers into $[0, 1]$; 2) the event that a pair of nodes are directly connected is independent of the event that another pair of nodes are directly connected.

Define $\mathrm{r}_l(k|x_0)$ to be the probability that two arbitrary nodes separated by Euclidean distance x_0 are k hops apart using GF in the Log-normal-Nakagami model. We start with k=1.

5.2 Probability of direct connection

Under the Log-normal-Nakagami model, two nodes separated by distance x are directly connected iff the RSS exceeds a given threshold $P_{\min}$. Without shadowing and small-scale fading, the model reduces to the unit disk model where $P_{\min} = CP_t r_0^{-\eta}$. With shadowing and fading, we have:

$$\begin{aligned} \mathrm{r}(P_N(x) \geq P_{\min}) &= \mathrm{r}(CP_t x^{-\eta} 10^{Z/10}\Omega_0 \geq CP_t r_0^{-\eta}) \\ &= \mathrm{r}(Z \geq 10\eta\ \mathrm{o}\ _{10}(\frac{x}{r_0 \Omega_0^{1/\eta}})) \quad (22) \\ &= \mathrm{r}(x \leq r_0 \Omega_0^{\frac{1}{\eta}}\ \mathrm{e}\ \mathrm{p}(\frac{Z}{10\eta}\ \ 10)) \quad (23) \end{aligned}$$

Thus two nodes are directly connected if either of the following two conditions is satisfied: 1) Given the distance x and Nakagami fades value ω_0, two nodes are directly connected iff the (random) shadowing fades

$Z \geq 10\eta$ o $_{10}(\frac{x}{r_0\omega_0^{1/\eta}})$. 2) Given that the shadowing fades z and Nakagami fades value ω_0, two nodes are directly connected iff their distance $x \leq r_0\omega_0^{1/\eta}$ e p$(\frac{z \text{ n } 10}{10\eta})$.

Based on the first condition, the probability of having a direct connection between two arbitrary nodes separated by x_0 is:

$$\text{r}_l(k=1|x_0) = \int_0^\infty \int_{10\eta \quad _{10}\frac{x}{r_0\omega_0^{1/\eta}}}^\infty q(z)dz\zeta_0(\omega_0)d\omega_0 \quad (24)$$

$$= \int_0^\infty \frac{1}{2}\left(1 - erf(\frac{10\eta \text{ o } _{10}(\frac{x}{r_0\omega_0^{1/\eta}})}{\sqrt{2\sigma^2}})\right)\zeta_0(\omega_0)d\omega_0 \quad (25)$$

where $erf(.)$ is the error function.

Remark 2: Without small-scale fading, viz. considering the log-normal shadowing model only, the probability of having a direct connection between two arbitrary nodes separated by x_0 is: $\frac{1}{2}(1-erf(\frac{10\eta \quad _{10}\frac{x}{r_0}}{\sqrt{2\sigma^2}}))$. Similarly, the following analysis can be reduced to the analysis without small-scale fading by simply removing the integral with respect to ω_0.

In order to derive $\text{r}_l(k|x_0)$ for $k > 1$, we use the second condition to study the probability of a direct connection. Define $r_N(z_S, \omega_S)$ to be the transmission range of a transmitter (S) conditioned on the shadowing fades and Nakagami fades being z_S and ω_S respectively. Then:

$$r_N(z_S, \omega_S) = r_0\omega_S^{1/\eta} \text{ e p}(\frac{z_S \quad 10}{10\eta}) \quad (26)$$

Therefore any node, whose RSS from the transmitter (S) has shadowing fades $Z_S \in [z_S, z_S + dz_S]$ and Nakagami fades $\Omega_S \in [\omega_S, \omega_S + d\omega_S]$, is directly connected to S iff its Euclidean distance to the transmitter is smaller than or equal to $r_N(z_S, \omega_S)$. This allows us to apply the analysis used in the unit disk model.

5.3 Distribution of the remaining distance

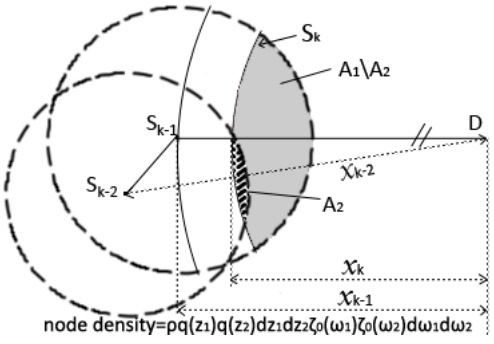


Fig. 4: Possible positions for the k^{th} hop node (S_k), are located on the arc. Consider the nodes whose RSS from S_{k-1} has fades $Z_1 \in [z_1, z_1 + dz_1]$ and $\Omega_1 \in [\omega_1, \omega_1 + d\omega_1]$; while its RSS from S_{k-2} has fades $Z_2 \in [z_2, z_2 + dz_2]$ and $\Omega_2 \in [\omega_2, \omega_2 + d\omega_2]$. The dashed-line circles represent the transmission range of S_{k-1} (resp. S_{k-2}) conditioned on the above values of shadowing and Nakagami fades. A_1 and A_2 are described in the following.

Define area size $A_1 = A(x_{k-1}, r_N(z_1, \omega_1), x_k)$ and $A_2 = A(x_{k-2}, r_N(z_2, \omega_2), x_k)$, where $A(x, r_1, r_2)$ and x_k are defined in Section 3. Define $f_l(x_k, k|x_0)$, $g_l(x_k, k|x_{k-1}, x_{k-2}, k-1)$, the event $\mathcal{B}_l$, $\text{r}_l(X_k \leq x_k, k|x_{k-1}, x_{k-2}, k-1)$ and $h_l(x_k, x_{k-1}, k|x_0)$ analogously as in Section 3 and use the subscript l to mark the corresponding probabilities in the Log-normal-Nakagami model. We will derive $\text{r}_l(X_k \leq x_k, k|x_{k-1}, x_{k-2}, k-1)$ by studying the following two events. Denote by $\mathcal{C}$ the event that there is at least one node whose Euclidean distance to the destination is smaller than x_k *and* has a direct connection to S_{k-1} *and* has no direct connection to S_{k-m} for $m \in [2, k]$ where S_0 is the source node. Denote by $\mathcal{D}$ the event that the node S_{k-1} is not directly connected to the destination. Events $\mathcal{C}$ and $\mathcal{D}$ are independent because of the independence of the shadowing and Nakagami fades. It is evident that:

$$\text{r}_l(X_k \leq x_k, k|x_{k-1}, x_{k-2}, k-1) = \text{r}(\mathcal{C}|\mathcal{B}_l) \times \text{r}(\mathcal{D}|\mathcal{B}_l) \quad (27)$$

We start with the analysis of event $\mathcal{C}$. In this paragraph we only consider the subset of nodes whose RSS from S_{k-1} has fades $Z_1 \in [z_1, z_1 + dz_1]$ and $\Omega_1 \in [\omega_1, \omega_1 + d\omega_1]$; while its RSS from S_{k-2} has fades $Z_2 \in [z_2, z_2 + dz_2]$ and $\Omega_2 \in [\omega_2, \omega_2 + d\omega_2]$. Due to the independence of the fades and the property of Poisson process, these nodes are distributed following a homogeneous Poisson process with intensity $\rho q(z_1)q(z_2)dz_1dz_2\zeta_0(\omega_1)\zeta_0(\omega_2)d\omega_1d\omega_2$. Denote by $\mathcal{E}$ the event that $Z_1 \in [z_1, z_1 + dz_1]$ *and* $Z_2 \in [z_2, z_2 + dz_2]$ *and* $\Omega_1 \in [\omega_1, \omega_1 + d\omega_1]$ *and* $\Omega_2 \in [\omega_2, \omega_2 + d\omega_2]$. $\text{r}(\mathcal{C}, \mathcal{E}|\mathcal{B}_l)$ is equal to the probability that there is at least one node in area $A_1 \setminus A_2$, as shown in Fig. 4. We approximate the size of area $A_1 \setminus A_2$ by $(A_1 - A_2)^+$, where $(A_1 - A_2)^+ = \text{ma } \{0, A_1 - A_2\}$. (Through this approximation we ignored some rare events that cause $A_1 - A_2 < 0$, which can possibly occur when $r_N(z_2, \omega_2)$ is much larger than $r_N(z_1, \omega_1)$. In contrast under the unit disk model it is always the case that $A_1 - A_2 \geq 0$.) Considering this subset of nodes only, $1 - \text{r}(\mathcal{C}, \mathcal{E}|\mathcal{B}_l)$ is equal to $1 - \text{e p}(-(A_1 - A_2)^+\rho q(z_1)q(z_2)dz_1dz_2\zeta_0(\omega_1)\zeta_0(\omega_2)d\omega_1d\omega_2)$, which is the probability that there is no node in area $A_1 \setminus A_2$. Note that A_1 depends on z_1 and ω_1; while A_2 depends on z_2 and ω_2.

Then considering all subset of nodes, we have:

$$\begin{aligned} &\text{r}(\mathcal{C}|\mathcal{B}_l) \\ &= 1 - \prod_{z_1,z_2 \in -\infty,+\infty \ ,\omega_1,\omega_2 \in 0,+\infty} \text{e p}(-(A_1 - A_2)^+\rho \\ &\quad q(z_1)q(z_2)dz_1dz_2\zeta_0(\omega_1)\zeta_0(\omega_2)d\omega_1d\omega_2 \quad (28) \\ &= 1 - \text{e p}(-\int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty (A_1 - A_2)^+\rho \\ &\quad q(z_1)q(z_2)dz_1dz_2\zeta_0(\omega_1)\zeta_0(\omega_2)d\omega_1d\omega_2) \quad (29) \end{aligned}$$

Since the event $\mathcal{D}$ only depends on x_{k-1}, we have:

$$\text{r}(\mathcal{D}|\mathcal{B}_l) = 1 - \text{r}_l(1|x_{k-1}) \quad (30)$$

Then substitute Eq. 29 and Eq. 30 into Eq. 27:

$$\begin{aligned} & \mathrm{r}_l(X_k \le x_k, k | x_{k-1}, x_{k-2}, k-1) \\ = \; & (1 - \mathrm{e}\ \mathrm{p}(-\int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty (A_1 - A_2)^+ \rho \\ & q(z_1) q(z_2) dz_1 dz_2 \zeta_0(\omega_1) \zeta_0(\omega_2) d\omega_1 d\omega_2)) \\ & \times (1 - \ \mathrm{r}_l(1 | x_{k-1})) \end{aligned} \tag{31}$$

By Leibniz integral rule:

$$\begin{aligned} & g_l(x_k, k | x_{k-1}, x_{k-2}, k-1) \\ = & \frac{\partial \ \mathrm{r}_l(X_k \le x_k, k | x_{k-1}, x_{k-2}, k-1)}{\partial x_k} \\ = & \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{\partial (A_1 - A_2)^+}{\partial x_k} \rho q(z_1) q(z_2) dz_1 dz_2 \\ & \zeta_0(\omega_1) \zeta_0(\omega_2) d\omega_1 d\omega_2 \\ & \times \mathrm{e}\ \mathrm{p}(-\int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty (A_1 - A_2)^+ \rho q(z_1) q(z_2) \\ & \zeta_0(\omega_1) \zeta_0(\omega_2) dz_1 dz_2 d\omega_1 d\omega_2)(1 - \ \mathrm{r}_l(1 | x_{k-1})) \end{aligned} \tag{32}$$

where $\partial A_1 / \partial x_k$ and $\partial A_2 / \partial x_k$ can be calculated by Eq. 6.

It is straightforward that for $k = 1$, we have:

$$\begin{aligned} & f_l(x_1, 1 | x_0) \\ = & \int_0^\infty \int_{-\infty}^\infty \frac{\partial A(x_0, r_N(z_1), x_1)}{\partial x_1} \rho q(z_1) dz_1 \zeta_0(\omega_1) d\omega_1 \\ & \times \mathrm{e}\ \mathrm{p}(-\int_0^\infty \int_{-\infty}^\infty A(x_0, r_N(z_1), x_1) \rho q(z_1) dz_1 \\ & \zeta_0(\omega_1) d\omega_1)(1 - \ \mathrm{r}_l(1 | x_0)) \end{aligned} \tag{33}$$

For $k = 2$, the pdf of the remaining distance of the previous hop is given by Eq. 33. Therefore:

$$h_l(x_2, x_1, 2 | x_0) = g_l(x_2, 2 | x_1, x_0, 1) f_l(x_1, 1 | x_0) \tag{34}$$

For $k > 2$, the joint pdf of x_k and x_{k-1} is calculated recursively:

$$\begin{aligned} h_l(x_k, x_{k-1}, k | x_0) \ = \ & \int_0^{x_0} g_l(x_k, k | x_{k-1}, x_{k-2}, k-1) \\ & h_l(x_{k-1}, x_{k-2}, k-1 | x_0) dx_{k-2} \end{aligned} \tag{35}$$

Finally for $k \ge 2$, we have:

$$f_l(x_k, k | x_0) = \int_0^{x_0} h_l(x_k, x_{k-1}, k | x_0) dx_{k-1} \tag{36}$$

5.4 Hop count statistics

Because of shadowing and small-scale fading, the destination can be possibly reached in a single hop no matter how far the remaining distance from that hop is. Therefore for $k \ge 2$:

$$\mathrm{r}_l(k | x_0) = \int_0^{x_0} \ \mathrm{r}_l(1 | x_{k-1}) f_l(x_{k-1}, k-1 | x_0) dx_{k-1} \tag{37}$$

Remark 3: Based on the above results, we can calculate the average number of hops for successful and unsuccessful transmissions, the probability of successful transmissions and the effective energy consumption by the same technique used in the unit disk model.

6 Simulation results

In this section, we report on simulations to validate the accuracy of the analytical results. The simulations are conducted in a wireless multi-hop network simulator written in C++. Nodes are deployed in a 400×400 square following a homogeneous Poisson process with intensity $\rho = 0.003$. The boundary effect is included in the simulation but it is shown to have a limited impact on the results. The route between two nodes is determined by the basic GF algorithm. The transmission range r_0 is varied from 10 to 50, which results in the average node degree varying from around 1 to 24. Note that r_0 is the transmission range without shadowing and small-scale fading. The value of r_0 can be specified by the network designer via adjusting the transmission power and receiver gain. The existence of a direct wireless link between an arbitrary pair of nodes will be further affected by shadowing and small-scale fading. Several values of the standard deviation in log-normal shadowing model have been used in our simulations, but only the results for $\sigma = 4$ are shown in this paper because other results show a similar trend. Further, we only include the results for $C_2 = 0.01$ and $Eng_c = 0.02$ (in Eq. 2) as an example and the value of Eng_c is found to have very limited impact on the results. In order to distinguish the impact on the network performance of different parameters, the packet error rate is not included (i.e. set $\alpha = 0$) except Fig. 8 and the small-scale fading is not included expect Fig. 6 and Fig. 7(b). Every point shown in the simulation result is the average value from 3000 simulations. As the number of instances of random networks used in the simulation is large, the confidence interval is too small to be distinguishable and hence is ignored in the following plots.

6.1 Hop count statistics

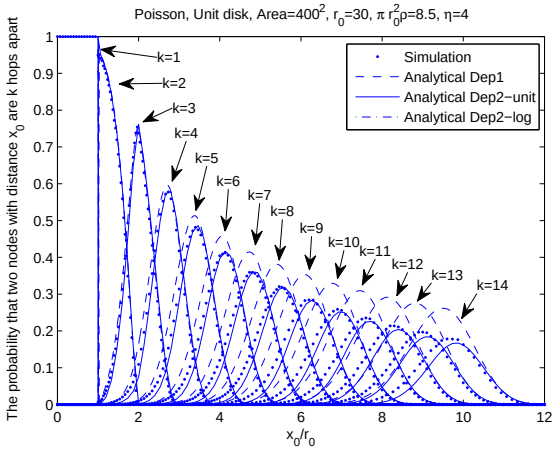
Fig. 5 shows the probability that two arbitrary nodes separated by distance x_0 are k hops apart using GF in the unit disk model and the log-normal shadowing model respectively. In the log-normal shadowing model [18], the received signal strength (RSS) attenuation (in dB) follows a normal distribution with standard deviation σ around the mean value. The mean value is given by the RSS under the path loss attenuation model, which is the model adopted in the unit disk model to determine the transmission range. Therefore when $\sigma = 0$, the log-normal shadowing model reduces to the unit disk model. As shown in Fig. 5, Dep2-unit and Dep2-log completely agree and the analytical results have a good match with the simulation results, which verifies the accuracy of our analysis in both the unit disk model and the log-normal shadowing model.

In addition, we can see that the accuracy is significantly improved by considering two previous hops (the result from this paper) compared with previous analysis considering only one previous hop (e.g. [23]). Further, it can be seen in Fig. 5 that the improvement of accuracy

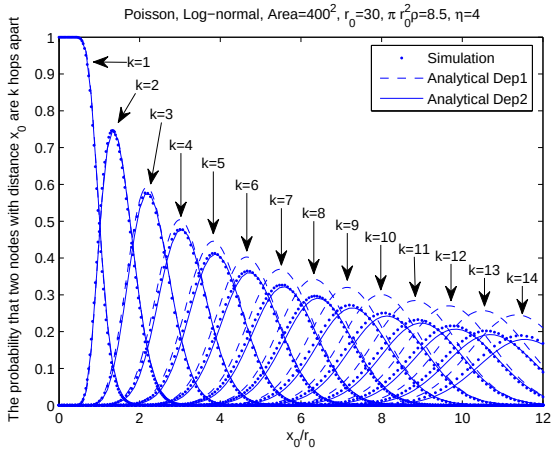
will be marginal if more than two previous hops are considered, which also confirms the analysis in Section 4. We expect this observation to be extended to many other areas (e.g. routing, localization, network security) and the approach used for shedding the independence assumption can be seen in a broader context. Specifically, our approach for shedding the independence assumption is to show that one can improve the accuracy of $\mathrm{r}(k|x)$ by taking into account the locations of previous m-hops nodes $(1 \leq m \leq k-1)$. However the improvement becomes marginal as $m > 2$. This suggests that the locations of nodes three or more hops away provide little information in determining the geometric relationship of a node with other nodes in the network. This observation further confirms our assertion in Section 4.

Furthermore, it is interesting to see that packets can be transmitted to a larger distance under the log-normal shadowing model than under the unit disk model, at the same number of hops. This is because log-normal shadowing introduces a Gaussian variation of the transmission range around the mean value, and with a higher chance a node can find a next-hop neighbor closer to the destination. This phenomenon is also observed in the study of connectivity [19].

Fig. 6 shows the probability that two arbitrary nodes separated by distance x_0 are k hops apart in the Log-normal-Nakagami model when the Nakagami parameter $m = 1$. Therefore, the corresponding network subjects to log-normal shadowing and Rayleigh fading. The result shown in Fig. 6 verifies the accuracy of our analysis. Further, it can be seen that Rayleigh fading reduces the probability that two nodes are connected by a path with k hops. This can be explained by the exponentially distributed RSS over the mean value caused by the Rayleigh fading which reduces the probability of direct connection. Therefore, Rayleigh fading has a negative impact on the network connectivity. A similar result is also observed in the next subsection.



(a) In the unit disk model, Dep2-unit is the result in the unit disk model from this paper, while Dep2-log is the result in log-normal shadowing model by letting $\sigma = 0$. Dep2-log is indistinguishable in the plot because the curve fully agrees with Dep2-unit.



(b) In the log-normal shadowing model

Fig. 5: The probability that two arbitrary nodes separated by Euclidean distance x_0 are k hops apart. $Dep1$ stands for the result calculated by considering the dependency on previous one hop. $Dep2$ is the result from this paper.

6.2 Effective energy consumption

Fig. 7 shows the probability of successful transmissions and the Eng_{eff}. It can be seen that, unsurprisingly, the probability of successful transmissions increases from nearly 0 to nearly 1 as r_0 increased from 10 to 50. In contrast, the effective energy consumption could hardly have been predicted by heuristic reasoning, and needs more explanations.

Take the results under the unit disk model as example. When r_0 is small, the network is made up of a large number of small components. An increase in r_0 will cause an increase in the size (number of nodes) of the components and also a reduction in the number of components. Therefore the average number of hops for unsuccessful transmission increases, and the energy wasted on unsuccessful transmission also increases. Thus there is an initial increase in Eng_{eff} with the increase in r_0. As r_0 further increases, although the average number of hops for successful/unsuccessful transmission still increases, the energy wasted on unsuccessful transmission starts to decrease as more source-destination pairs become connected. The balance of the two effects causes Eng_{eff} to peak at $r_0 \approx 19$. Above this transmission range, the decrease in wasted energy starts to dominate, which causes a subsequent decrease in Eng_{eff}. As r_0 increases further, the average number of hops approaches its maximum and the energy wasted on unsuccessful transmission also reduces to a small amount. These cause Eng_{eff} to reach its minimum at $r_0 \approx 31$. Above this transmission range, most source-destination pairs are connected as shown in Fig. 7 (a.1). Another effect starts to dominate. That is, the increase in r_0 causes the increase in the per-hop energy consumption (like r_0^2) and the decrease in the number of hops (approximately like $1/r_0$). The net effect is an increase in Eng_{eff} with the increased r_0. Most previous studies have only considered this last stage of the relation between the energy consumption

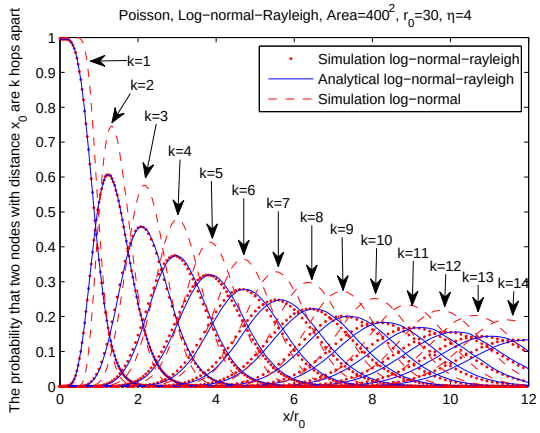
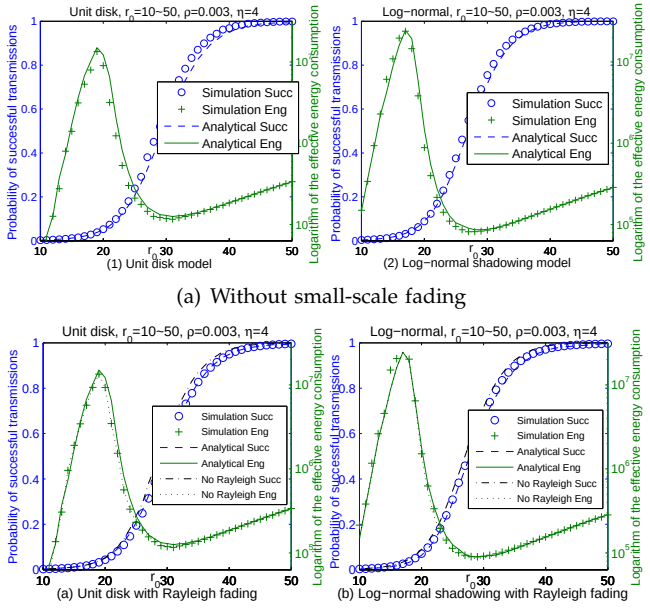


Fig. 6: The probability that two arbitrary nodes separated by distance x_0 are k hops apart in a network subject to log-normal shadowing and Rayleigh fading.



(a) Without small-scale fading

(b) With small-scale fading

Fig. 7: Probability of successful transmissions (Succ) and effective energy consumption (Eng) in the network. Subfigure (a) shows the results without small-scale fading and (b) shows the results with Rayleigh fading. Further, "No Rayleigh" is the result without small-scale fading shown in (b) for comparison.

and the transmission range and therefore cannot give a complete understanding of the energy efficiency in end-to-end packet transmissions.

It is interesting to note that the energy optimizing transmission range is around 31, which corresponds to a network with most (around 70%) source-destination pairs connected but not all of them. (Note that for $r_0 < 20$, Eng_{eff} may be smaller than the minimum Eng_{eff}. However at such value of r_0 most source-destination pairs are disconnected using GF and no meaningful service can be provided by the wireless multi-hop network.) In order for more than 99% source-destination pairs to be connected, r_0 has to be larger than 47 and Eng_{eff} will increase to more than 225% of its minimum value in the unit disk model. A similar result can also be found in the log-normal shadowing model and the models with Rayleigh fading. Therefore significant energy savings can be obtained by requiring most nodes, instead of all nodes, in the network to be connected. This observation also agrees with the analytical results in [12]. In addition, our result gives the amount of energy that can be saved. Purely from an energy-saving perspective and without consideration of other implications, this interesting result shows that the most energy-efficient topology control algorithms should be designed to let 70% (under this network setting) of the source-destination pairs be connected at the same time. The result sheds insight on the design of large wireless multi-hop networks where energy-efficiency is a important issue.

Further, Fig. 7 (b) shows that the probability of successful transmissions is slightly lower in a network with Rayleigh fading compared to a network without Rayleigh fading. This confirms our assertion in the previous subsection.

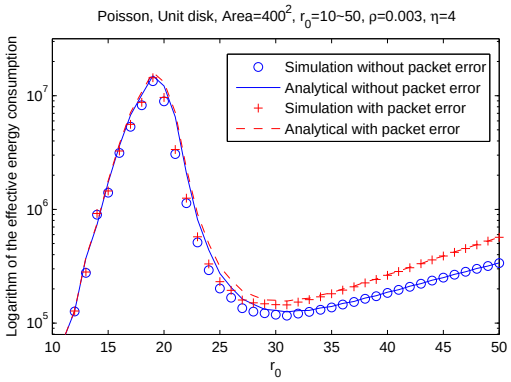


Fig. 8: The effective energy consumption subject to packet error.

Fig. 8 shows the effective energy consumption with a non-zero packet error rate as shown in Eq. 2. The packet error rate increases from 0.004 when $r_0 = 10$ to 0.40 when $r_0 = 50$. It can be seen that as the transmission range increases, the tail of the effective energy consumption increases faster than its error-free counterpart. This is because an increase in the transmission range causes an increase in the number of neighbors and also an increase in the distance between the transmitter and the receiver. This in turn increases the packet error rate and the energy consumption. Therefore when the packet error rate is non-zero, the energy optimizing transmission range becomes smaller as can be seen in Fig. 8.

Fig. 9 shows the effective energy consumption under the log-normal shadowing model with various values of standard deviations. It can be seen that a larger variance in the log-normal shadowing model leads to a lower energy consumption and a smaller optimum transmission range. This is because a larger variance provides a larger probability for a node to forward the packet to a further node that is closer to the intended destination, which is similar to the observation obtained in Section 6.1.

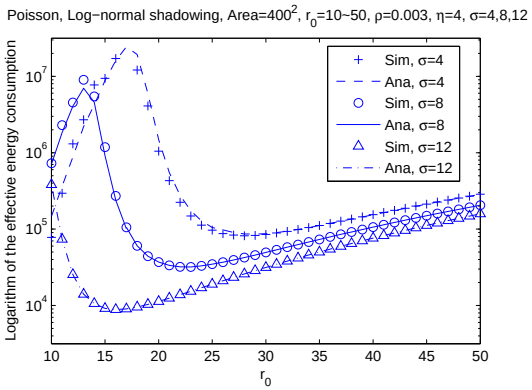


Fig. 9: The effective energy consumption under the log-normal shadowing model with various values of standard deviations.

6.3 Impact of node density and path loss exponent on the optimum transmission range

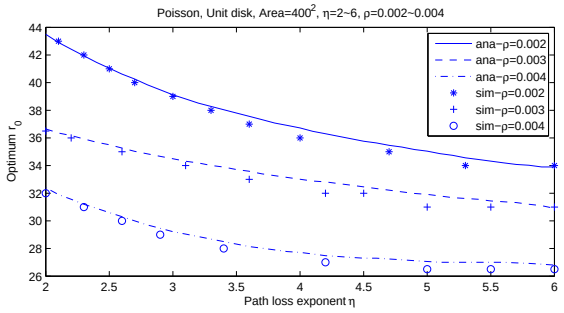


Fig. 10: Impact of node density and path loss exponent on the optimum transmission range.

Fig. 10 illustrates the impact of node density and path loss exponent on the optimum transmission range in the unit disk model. It can be seen that an increase in the node density will cause a decrease in the optimum transmission range. It is because an increase in node density without reduction in transmission range causes an increase in the average number of neighbors as well as the probability of successful transmissions between two nodes. Fig. 10 also shows that a higher path loss exponent will result in a smaller optimum transmission range. This is because an increase in the path loss exponent will cause an increase in the per-hop energy consumption, as given by Eq. 2. Therefore under a higher value of the path loss exponent it is more energy-efficient to have smaller components, hence a smaller optimum transmission range.

It has been shown that the probability $\mathrm{r}(k|x)$ and the energy consumption are affected by the node density and path loss exponent. Our analysis fully captures these effects and sheds insight on the design of a wireless multi-hop network.

7 Conclusions and Future Work

We investigated the hop count statistics and the energy consumed in the end-to-end packet transmissions in a wireless multi-hop network. Considering both shadowing and small-scale fading, we obtained analytical results on the probability distribution of the number of hops between two arbitrary nodes. Further, we analyzed the impact of the spatial dependence problem on the $\mathrm{r}(k|x)$. Considering the randomness of node deployment and a complex radio environment which may result in disconnected paths between nodes, we derived the distribution of the number of hops traversed by packets before being dropped if the transmission is unsuccessful. As an application of the above results, we derived the effective energy consumption per successfully transmitted packet in end-to-end packet transmission. We showed that there exists an optimum transmission range which minimizes the effective energy consumption. The research provides useful guidelines on the design of a multi-hop network in the presence of shadowing and fading.

The hop count statistics obtained in this paper will also be useful to determine other aspects of wireless multi-hop network performance, e.g. end-to-end throughput and delay. Allowing disconnected paths enables our results to be applicable to sparse network, which is essential to the study of partial connectivity [38]. Moreover, we plan to study the hop count statistics and energy consumption in a mobile ad-hoc network in the future.

References

[1] W. Hu, V. N. Tran, N. Bulusu, C. T. Chou, S. Jha, and A. Taylor, "The design and evaluation of a hybrid sensor network for cane-toad monitoring," in *Fourth International Symposium on Information Processing in Sensor Networks*, 2005, pp. 503– 508.

[2] M. Conti and S. Giordano, "Multihop ad hoc networking: The reality," *IEEE Communications Magazine*, vol. 45, pp. 88–95, 2007.

[3] X. Ta, G. Mao, and B. D. O. Anderson, "On the probability of k-hop connection in homogeneous wireless sensor networks," *IEEE Communication Letters*, vol. 11, no. 8, pp. 662–664, 2007.

[4] ——, "Evaluation of the probability of k-hop connection in homogeneous wireless sensor networks," in *IEEE Global Telecommunications Conference*, Washington DC, USA, 2007, pp. 1279–1284.

[5] J.-C. Kuo and W. Liao, "Hop count distribution of multihop paths in wireless networks with arbitrary node density: Modeling and its applications," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, 2007.

[6] S. Dulman, M. Rossi, P. Havinga, and M. Zorzi, "On the hop count statistics for randomly deployed wireless sensor networks," *International Journal of Sensor Networks*, vol. 1, pp. 89–102, 2006.

[7] M. Zorzi and R. R. Rao, "Geographic random forwarding (geraf) for ad hoc and sensor networks: Multihop performance," *IEEE Transactions on Mobile Computing*, vol. 2, no. 4, pp. 337–348, 2003.

[8] S. De, A. Caruso, T. Chaira, and S. Chessa, "Bounds on hop distance in greedy routing approach in wireless ad hoc networks," *International Journal of Wireless and Mobile Computing*, vol. 1, no. 2, pp. 131 – 140, 2006.

[9] F. Xue and P. R. Kumar, "The number of neighbors needed for connectivity of wireless networks," *Wireless Networks*, vol. 10, no. 2, pp. 169–181, 2004.

[10] P. Padhy, K. Martinez, A. Riddoch, H. L. . R. Ong, and J. K. Hart, "Glacial environment monitoring using sensor networks," in *Real-World Wireless Sensor Networks Workshop*, 2005, pp. 10–14.

[11] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[12] X. Ta, G. Mao, and B. D. O. Anderson, "On the properties of giant component in wireless multi-hop networks," in *IEEE Infocom*, 2009.

[13] S. A. G. Chandler, "Calculation of number of hops required in randomly located radio network," *Electronics Letters*, vol. 25, pp. 1669–1671, 1989.

[14] M. Mauve, J. Widmer, and H. Hartenstein, "Survey on position-based routing in mobile ad hoc networks," *IEEE Networks*, vol. 15, no. 6, pp. 30–39, 2001.

[15] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad Hoc Networks*, vol. 3, no. 3, pp. 325 – 349, 2005.

[16] P. Acevedo Contla and M. Stojmenovic, "Estimating hop counts in position based routing schemes for ad hoc networks," *Telecommunication Systems*, vol. 22, no. 1, pp. 109–118, 2003.

[17] Z. Zhang, G. Mao, and B. D. Anderson, "On the effective energy consumption in wireless sensor networks," in *Proceedings IEEE WCNC*, 2010, pp. 1 – 6.

[18] T. S. Rappaport, *Wireless Communications: Principles and Practice, 2nd Edition*. Prentice Hall, 2001.

[19] R. Hekmat and P. V. Mieghem, "Connectivity in wireless ad-hoc networks with a log-normal radio model," *Mobile Networks and Applications*, vol. 11, no. 3, pp. 351–360, 2006.

[20] S. Mukherjee and D. Avidor, "Connectivity and transmit-energy considerations between any pair of nodes in a wireless ad hoc network subject to fading," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 2, pp. 1226–1242, 2008.

[21] D. Miorandi and E. Altman, "Coverage and connectivity of ad hoc networks in presence of channel randomness," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 1, 2005, pp. 491– 502.

[22] M. Haenggi, "On routing in random rayleigh fading networks," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1553 – 1562, 2005.

[23] J. Deng, Y. S. Han, P.-N. Chen, and P. K. Varshney, "Optimal transmission range for wireless ad hoc networks based on energy efficiency," *IEEE Transactions on Communications*, vol. 55, no. 9, pp. 1772–1782, 2007.

[24] R. Zhang and J.-M. Gorce, "Optimal transmission range for minimum energy consumption in wireless sensor networks," in *IEEE Wireless Communications and Networking Conference*, 2008, pp. 757 – 762.

[25] C. Bettstetter and C. Hartmann, "Connectivity of wireless multihop networks in a shadow fading environment," *Wireless Networks*, vol. 11, no. 5, pp. 571 – 579, 2005.

[26] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels*. John Wiley and Sons, Inc., 2000.

[27] M. M. Carvalho and J. J. Garcia-Luna-Aceves, "Delay analysis of IEEE 802.11 in single-hop networks," in *IEEE International Conference on Network Protocols*, 2003, pp. 146–155.

[28] O. Kasten, "Energy consumption," 2001. [Online]. Available: http://www.inf.ethz.ch/personal/kasten/research/bathtub/

[29] G. Xing, C. Lu, R. Pless, and Q. Huang, "Impact of sensing coverage on greedy geographic routing algorithms," *IEEE Transactions on Parallel and Distributed Systems*, vol. 17, no. 4, pp. 348– 360, 2006.

[30] L. E. Miller, "Distribution of link distances in a wireless network," *Journal of Research of the National Institute of Standards and Technology*, vol. 106, pp. 401–412, 2001.

[31] E. W. Weisstein, "Circle-circle intersection," 2009. [Online]. Available: http://mathworld.wolfram.com/Circle-CircleIntersection.html

[32] S. Dulman and P. Havinga, "Statistically enhanced localization schemes for randomly deployed wireless sensor networks," in *Proceedings of the Intelligent Sensors, Sensor Networks and Information Processing Conference*, 2004, pp. 403–410.

[33] N. Lo and H.-Y. Kuo, "Two-hops neighbor-aware routing protocol in mobile ad hoc networks," in *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2008, vol. 4611/2007, pp. 340–349.

[34] C. J. Mallery, S. Medidi, and M. Medidi, "Relative localization with 2-hop neighborhood," in *International Symposium on a World of Wireless, Mobile and Multimedia Networks*, 2008, pp. 1 – 4.

[35] I. Khalil, S. Bagchi, and N. Shroff, "Liteworp: A lightweight countermeasure for the wormhole attack in multihop wireless networks," in *Proceedings of the 2005 International Conference on Dependable Systems and Networks*, 2005, pp. 612 – 621.

[36] R. Nelson, *Probability, stochastic processes, and queueing theory: the mathematics of computer performance modelling*. Springer, 1995.

[37] M. Franceschetti and R. Meester, *Random networks for communication: From Statistical Physics to Information Systems*. Cambridge University Press, 2007.

[38] Z. Zhang, S. C. Ng, G. Mao, and B. D. O. Anderson, "On the k-hop partial connectivity in finite wireless multi-hop networks," in *Proceedings IEEE ICC*, 2011, pp. 1–5.

[39] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming, Eds. W. M. McEneany, G. Yin, Q. Zhang.*, p. 547 566, 1998.

[40] G. Mao, "Research on wireless multi-hop networks: Current state and challenges," in *submitted as invited position paper to International Conference on Computing, Networking and Communications*, 2012.

[41] S. C. Ng, G. Mao, and B. D. Anderson, "Analytical bounds on the critical density for percolation in wireless multi-hop networks," in *accepted by GLOBECOM*, 2011.

[42] S. C. Ng, W. Zhang, Y. Zhang, Y. Yang, and G. Mao, "Analysis of access and connectivity probabilities in vehicular relay networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 1, p. 140, 2011.

[43] S. Rajabi, M. Shahabadi, and M. ArdebiliPoor, "Modeling of the correlation coefficients of a receive antenna array in a mimo multipath channel," in *IEEE/IFIP International Conference in Central Asia on Internet*, 2006, pp. 1–4.

[44] P. Agrawal and N. Patwari, "Correlated link shadow fading in multi-hop wireless networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, p. 40244036, 2009.

[45] C. X. Wang, M. Patzold, and Q. Yao, "Stochastic modeling and simulation of frequency-correlated wideband fading channels," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 3, pp. 1050 – 1063, 2007.

Zijie Zhang received his BEng degree in Electronic and Communications Engineering from the University of Hong Kong in 2008. He is currently a PhD candidate in the School of Electrical and Information Engineering at the University of Sydney. His research interests include wireless multi-hop networks, vehicular ad-hoc networks, mobile ad-hoc networks and graph theory.

Guoqiang Mao received PhD in telecommunications engineering in 2002 from Edith Cowan University. He joined the School of Electrical and Information Engineering, the University of Sydney in December 2002 where he is a Senior Lecturer now. His research interests include wireless localization techniques, wireless multi-hop networks, graph theory and its application in networking, and network performance analysis. He is a Senior Member of IEEE.

Brian D. O. Anderson (M'66-SM'74-F'75-LF'07) was born in Sydney, Australia, and educated at Sydney University in mathematics and electrical engineering, with PhD in electrical engineering from Stanford University in 1966. He is a Distinguished Professor at the Australian National University and Distinguished Researcher in National ICT Australia. His awards include the IEEE Control Systems Award of 1997, the 2001 IEEE James H Mulligan, Jr Education Medal, and the Bode Prize of the IEEE Control System Society in 1992, as well as several IEEE best paper prizes. He is a Fellow of the Australian Academy of Science, the Australian Academy of Technological Sciences and Engineering, the Royal Society, and a foreign associate of the National Academy of Engineering. His current research interests are in distributed control, sensor networks and econometric modelling.

SUPPLEMENTARY MATERIAL

8 RELATED WORK

The hop count statistics were first investigated by Chandler [13] in 1989. He analyzed the probability that two randomly chosen nodes separated by a known distance can communicate in k or less hops where nodes are uniformly distributed over a plane. However the analysis was incomplete as the aforementioned spatial dependence problem was incorrectly ignored. Ta et al. [3] investigated the probability $\mathrm{r}(k|x)$ for nodes Poissonly distributed in a square. They pointed out the spatial dependence problem in the analysis of the probability $\mathrm{r}(k|x)$. Later in [4] the same authors empirically improved their earlier result in [3] by considering the impact of boundary effect and the spatial dependence problem.

In real applications, the packets are forwarded from a source to a destination according to certain routing algorithms. Many routing algorithms (e.g. LEACH, AODV or geographic routing [14]) share the similar idea of greedy forwarding (GF) [14], that is, to forward the message to the node that is closest to the destination [15]. Much research on the hop count statistics is based on the distributed routing algorithm GF, however the spatial dependence problem were incorrectly ignored. Specifically, two nodes are k hops apart if the path between them, *using GF*, is k hops. Zorzi et al. [7] proposed a GF algorithm for a network where nodes are Poissonly distributed in the coverage area of a transmitting node. They studied an upper and a lower bound on the average number of hops between two nodes separated by a known Euclidean distance, where the focus of this paper is on a complete characterization of the probability distribution of the number of hops between two arbitrary nodes in the network. **In [16], Contla and Stojmenovic considered position based routing schemes for a wireless multi-hop network where nodes are uniformly distributed in a square. They studied the average number of hops between an arbitrary pair of source-destination nodes.** As pointed out in Section 1, many applications need the knowledge of the probability distribution of the number of hops, instead of just the mean value. Dulman et al. [6] investigated the probability $\mathrm{r}(k|x)$ by estimating the expected progress per hop using GF. They considered the impact of the Euclidean distance between neighboring nodes in the previous hop on the progress in the current hop. Both [7] and [6] were established on the assumption that a packet can always reach the destination using GF. Further, the aforementioned research only considered the impact of one previous hop in their studies. Recently, the accuracy of the probability $\mathrm{r}(k|x)$ was significantly improved in [17] by considering the spatial dependence of two-hop neighbors.

The aforementioned results are all based on the *unit disk communication model*, in which two nodes are directly connected if and only if (iff) the Euclidean distance between them is smaller than or equal to the transmission range. The unit disk model is simple but unrealistic [18]. Considering the log-normal shadowing model, Hekmat and Mieghem [19] showed, through simulations, that the probability of a network being connected increases with increasing value of the shadowing parameter, which is the ratio between the standard deviation of shadowing and the path loss exponent [18]. Mukherjee and Avidor [20] considered the impact of the log-normal shadowing on the probability $\mathrm{r}(k|x)$ in a wireless ad hoc network where nodes are Poissonly distributed in a disk, which ignored the spatial dependence problem. In addition to shadowing, the communication between two nodes can be affected by the small-scale fading (e.g. Rayleigh fading). From the connectivity point of view, Miorandi and Altman [21] studied the node isolation probability in a network subject to both log-normal shadowing and Rayleigh fading. They showed that Rayleigh fading reduces the connectivity probability of the network. Moreover, Haenggi [22] studied the routing performance for large multi-hop networks, considering the impact of Rayleigh fading on the end-to-end delivery probability. It is shown that routing over many short hops is not as beneficial in a network subject to Rayleigh fading as that for a network without Rayleigh fading. In this paper, we considered the impact of both log-normal shadowing and small-scale fading on the hop count statistics. Further, our analysis takes into account the impact of the spatial dependence problem, which is a major technical hurdle in the accurate analysis of the probability $\mathrm{r}(k|x)$.

The analysis on the hop count statistics can be used in a number of areas in wireless multi-hop networks. This paper focuses on its use in energy-efficient operations of wireless multi-hop networks as an example. Minimizing energy consumption is one of the major considerations in the design of battery powered wireless multi-hop networks. In many applications it is difficult to change or re-charge a battery for the wireless nodes. From a designer point of view, a popular approach of reducing energy consumption is optimally choosing the transmission power. In [23], Deng et al. considered a network where nodes are Poissonly distributed in a circular area. They assumed that there is always a path between any pair of nodes using GF. By analyzing the average progress per hop that a packet is transmitted towards the destination, they obtained analytical results on the distance-energy efficiency, which is the ratio of the average progress to the energy consumed in a single transmission, and the optimum transmission range that maximizes the distance-energy efficiency for high-density networks. Zhang and Gorce [24] considered the impact on energy consumption of unreliable links. They postulated that with a certain probability a transmission between two directly connected nodes is unsuccessful, re-transmissions may then be required and energy consumption may be consequently higher. The extra energy consumed due to unreliable links is also considered in

this paper.

9 Usefulness of the hop count statistics

The results on the hop count statistics hold the key to solving a large number of problems in wireless multi-hop networks:

- The summation of $\mathrm{r}(k)$ from $k = 1$ to $k = \infty$ provides the probability that two randomly chosen nodes are connected (via a multi-hop path), which in turn can lead to result on the probability of a connected network. The result will be valid for not only large-scale networks [39] but also for small-scale networks which are often encountered in real applications. So far, there are few analytical results on the connectivity of small-scale networks.
- The hop count statistics are also useful in network capacity analysis. In [40], it is demonstrated that the capacity scaling law of multi-hop networks observed in [11] can be easily explained by the increase in the average number of hops (hence the increase in the portion of bandwidth spent on relaying traffic) as the network becomes larger.
- The probability $\mathrm{r}(k)$ is also useful in estimating the energy consumption, the network lifetime and the reliability of end-to-end packet transmission [7]. As shown in this paper, $\mathrm{r}(k)$ can be used to estimate the effective energy consumption and help the network designers to choose the optimum transmission range/power to minimize the energy consumption. It can also be readily used to help a network designer to set the transmission range/power to provide a guaranteed performance on the end-to-end packet transmissions.
- The probability $\mathrm{r}(k|x)$ has been used in [41] to form a novel approach to obtain bounds on the critical density for percolation in wireless multi-hop networks, a well-known open problem in the area. In [38] results on $\mathrm{r}(k|x)$ are used as a main tool to study the partial connectivity of a wireless multi-hop network with infrastructure support. Besides its use in performance analysis, a protocol designer can use results on $\mathrm{r}(k|x)$ to help choose the optimum protocol parameters, e.g. the timeout parameter TTL used in many routing protocols, to balance bandwidth (or energy) consumption and probability of successful delivery [5].
- The probability density $\mathrm{r}(x|k)$ is useful in estimating the distance between two nodes from their neighborhood information and obtaining variance of such an estimate, which has in turn been used in forming a localization algorithm with improved performance [8], [32].
- Further, the technique used to derive $\mathrm{r}(k|x)$ can be simplified to study 1D networks or grid networks so that the analysis can be applied to study vehicular networks, see [42] for an example where $\mathrm{r}(k|x)$ is used to derive the access and connectivity probabilities of 1D vehicular networks.

10 Reliability of the assumption of the independence between links

In some environments, the assumption of independence of connections may not be accurate while in other environments (e.g. open space) it is a reasonable assumption. For example, it is generally accepted that if a pair of transmitters are separated by more than $\lambda/4$, where λ is the wavelength, their signals at a common receiver can be regarded as statistically independent. Further it was shown [43] that if a pair of receivers are separated by more than λ, their received signals from a common transmitter are only weakly correlated (with a correlation coefficient less than 0.15). At a typical frequency of 5GHz, $\lambda = 0.06$m. Thus the requirement on the separation of vehicles can be easily met. We also note that although field measurements in real applications seem to indicate that the connectivity between different pairs of geographically/frequency proximate wireless nodes are correlated [44], [45], the independence assumption is generally considered appropriate for far-field transmission and has been widely used in the literature under many channel models including the log-normal shadowing model [18], [20], [25].

A Robust Reachability Review for Control System Security

Adrian N. Bishop

Abstract

Control systems underpin the core technology in numerous critical infrastructure systems; e.g. the electricity, transportation and many defence systems. Increasingly, such systems are becoming the target of a novel type of deliberate cyber-attack. The notion of control system security is a modern idea concerned with the analysis, design and application of tools that ensure the operational goals of a particular control systems are protected from deliberate, malicious, electronic attack. The aim of this field is to reduce the likelihood of success, and the severity of impact, of a cyber-attack against control systems operating within critical infrastructure. This contribution re-examines a classical notion of control reachability through set-theoretic arguments with an additional, modern, emphasis on control system security. In particular, the reachability idea is extended to a compromised control system architecture. Classical and novel results on reachability are defined within this setting from the point-of-view of an attacker and, conversely, a system designer wanting to secure the system. The idea of reachability studied in this setting for secure control is important in both the design of robust control systems with security features and in assessing the vulnerability of particular control systems.

I. Introduction

The problem of networked estimation and control has been considered extensively in recent years; see the expositions in [1]–[3]. This work considers *security for networked estimation and control systems*; or so-called networked cyber-physical systems. There are many networked control systems that operate critical infrastructure within strategic resource sectors[1]. The reliability and security of most systems of this nature is critical. However, the distributed and networked nature of the system often makes it vulnerable to a deliberate attack; e.g. many distributed cyber-physical systems are controlled via unsecured communication networks that may also be accessible through the world-wide-web [4]. Moreover, there are many sensitive applications involving networked control systems, e.g. power and transportation networks, and such applications are appealing targets for many malicious entities. A disruption or disconnection of these services may have significant social, environmental and/or economic consequences. Secure networked estimation and control is thus an important emerging problem[2].

The idea of secure control and estimation systems is becoming increasingly important, e.g. as a result of significant publicised attacks and the realization that a compromised system operating in a key sector may have significant social, environmental and/or economic consequences. Early studies in [4], [8] outline a framework for discussing and investigating the security of networked control and estimation systems. In [4] two classes of attack are characterized: namely *denial-of-service (DOS) attacks* where the attacker prevents the transmission of, e.g., measurement or control signals, and; *deceptive attacks* where a false (or distorted) signal is inserted into the system to disrupt the behaviour of the estimator or controller.

The problem of robust control in the presence of DOS attacks has been investigated in [8]. This work has similarities with those networked control problems that are designed to cope with limited-data rate communications and packet-dropping. However, the latter work often assumes some random model of packet failure whereas any DOS attack may be far more systematic. The problem of secure control and estimation under false-data attacks has been investigated. In [9], [10] the idea of reachability is used to analyze the security of control systems and their resilience to attack. In [11] a necessary and sufficient condition is provided under which an attacker can destabilize a system nominally controlled by an optimal control law while remaining undetected by a particular detector.

Alternatively, the problem of secure estimation and control in networked systems under false-data attacks has attracted significant attention recently; e.g. see [11]–[27]. At a high-level, the majority of this work addresses the problem of state estimation with inherent detection of false-data attacks or, conversely, a characterization of those attacks which are likely (under some modelling and algorithmic assumptions) to remain undetected in certain state estimation schemes. The classical field of fault-detection and fault-tolerant control; see [28], is related to this work also.

This contribution examines a classical notion of control reachability through set-theoretic arguments with a modern emphasis on control system security. In particular, the reachability idea is extended to a compromised control system architecture. Classical and novel results on reachability are defined within this setting from the point-of-view of an attacker and, conversely, a system designer wanting to secure the system. The idea of reachability studied in this setting for secure control is important in both the design of robust control systems with security features and in assessing the vulnerability of particular control systems.

II. A General Introduction on Robust Reachability

The introduction in this section follows the classical results in [29]–[31]. However, the notation and set-theoretic reachability results are stated, and derived, in a different fashion that suits the later framework of control security.

Consider the time-varying discrete-time system

$$\mathbf{x}_{k+1} = f_k(\mathbf{x}_k, \mathbf{u}_k) + g_k(\mathbf{w}_k) \tag{1}$$

A.N. Bishop is with NICTA and the Australian National University.

[1]Examples of so-called networked cyber-physical systems include critical infrastructure such as transportation networks, power and electricity grids, water distribution networks, etc. In these cases, the networked components may be spatially separated.

[2]An example of an attack on a real-world networked control system is the highly sophisticated *Stuxnet* attack on the Iranian nuclear energy program [5]. Stuxnet is a worm that spreads indiscriminately across computer networks running Microsoft Windows causing no disruption but which also includes a payload designed to target Siemens supervisory control and data acquisition (SCADA) systems configured to control and monitor very specific industrial processes (most likely certain centrifuges). In particular, Stuxnet infects the programmable logic controllers (PLC) in a SCADA system by subverting the application used to reprogram these devices. It then disrupts the operation of certain, very specific, industrial processes being driven by the controller. Stuxnet is so sophisticated that the general consensus is that it was state-sponsored and the first such act of industrial cyber-warfare [5]. Other publicised, but less dramatic, examples of intrusions and attacks on networked control systems exist; e.g. [6], [7].

defined for $k \in \mathbb{N} \cup \{0\} = \mathbb{N}_0$ and where $\mathbf{x}_k \in \mathcal{X}_k \subseteq \mathcal{X} \subseteq \mathbb{R}^n$ is the *state*, $\mathbf{u}_k \in \mathcal{U}_k \subseteq \mathcal{U} \subseteq \mathbb{R}^m$ is the *control vector* and $\mathbf{w}_k \in \mathcal{W}_k \subseteq \mathcal{W} \subseteq \mathbb{R}^l$ is the *system uncertainty input* or error. The functions $f_k : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}^n$ and $g_k : \mathcal{W}_k \to \mathbb{R}^n$ are known and $f_k + g_k : \mathcal{W}_k \times \{\mathbb{R}^n \times \mathcal{U}\} \to \mathcal{X}_{k+1}$. The system (1) describes the physical system under consideration.

The following definition makes matters precise.

Definition 1. *Suppose $\mathcal{X}_0$ is given and fixed. Suppose a sequence of control functions are specified of the form $\mathbf{u}_k = u_k(\mathbf{x}_k) : \mathcal{X}_k \to \mathcal{U}_k$ where*

$$\mathcal{X}_{k+1} = \{\mathbf{x}_{k+1} \in \mathbb{R}^n : \mathbf{x}_{k+1} = f_k(\mathbf{x}_k, \mathbf{u}_k) + g_k(\mathbf{w}_k), \\ \forall \mathbf{x}_k \in \mathcal{X}_k,\ \forall \mathbf{u}_t \in \mathcal{U}_t,\ \forall \mathbf{w}_t \in \mathcal{W}_t,\ \forall t \in \{0, \ldots, k\}\}$$

*is then called the **local set of admissible state values under the sequence** $\{\mathcal{U}_t\}_{t=0}^k$ at time k for all $k > 0$.*

If we suppose that $\mathbf{u}_k$ take any value in $\mathcal{U}$ at time k and suppose $\mathcal{X}_0$ is given and fixed then

$$\mathfrak{X}_{k+1} = \{\mathbf{x}_{k+1} \in \mathbb{R}^n : \mathbf{x}_{k+1} = f_k(\mathbf{x}_k, \mathbf{u}_k) + g_k(\mathbf{w}_k), \\ \forall \mathbf{x}_k \in \mathfrak{X}_k,\ \forall \mathbf{u}_t \in \mathcal{U},\ \forall \mathbf{w}_t \in \mathcal{W}_t,\ \forall t \in \{0, \ldots, k\}\}$$

*is called the **global set of admissible state values** at time k for all $k > 0$.*

It is clear that $\mathcal{X}_k \subseteq \mathfrak{X}_k$. Note $\mathcal{X} = \cup_k \mathcal{X}_k \subseteq \mathbb{R}^n$ is defined by a particular feedback controller whereas $\mathcal{U}$ and $\mathcal{W}$ are defined by the physical properties of the system as expected. Obviously $\mathcal{X} \subseteq \cup_k \mathfrak{X}_k$.

The general control problem under consideration is stated as follows.

Problem 1. *Consider a time $0 < \tau \in \mathbb{N}$ and suppose $\mathcal{X}_0$ is given and fixed. Design a control function $\mathbf{u}_k = u_k(\mathbf{x}_k) : \mathcal{X}_k \to \mathcal{U}_k$ mapping $\mathbf{x}_k$ into $\mathcal{U}_k$, for all $k \in \{0, \ldots, \tau - 1\}$ such that the following holds*

$$f_{\tau-1} + g_{\tau-1} : \mathcal{W}_{\tau-1} \times \{\mathcal{X}_{\tau-1} \times \mathcal{U}_{\tau-1}\} \to \mathcal{X}_\tau^* \tag{2}$$

for all $\mathbf{w}_k \in \mathcal{W}_k$, $\forall k \in \{0, \ldots, \tau - 1\}$. The set $\mathcal{X}_\tau^* \in \mathbb{R}^n$, $0 < \tau \in \mathbb{N}$ constitutes the desired, or target, set at time τ for the controlled system (1) and should be known to the control designer a priori.

A question in robust control then concerns the solvability of Problem 1. We introduce the following definition.

Definition 2. *A set $\mathcal{X}_\tau^* \in \mathbb{R}^n$, $0 < \tau \in \mathbb{N}$ is said to be **reachable given** $\mathcal{X}_0$ if there exists, at least, a single sequence of control functions $u_k : \mathcal{X}_k \to \mathcal{U}_k$, $\forall k \in \{0, \ldots, \tau - 1\}$ that solves Problem 1.*

Obviously $\mathcal{X}_\tau^* \subseteq \mathfrak{X}_\tau$ is a necessary condition for reachability, however, it is not necessarily true that every subset of $\mathfrak{X}_\tau$ is a reachable subset given some $\mathcal{X}_0$. Admissible is not equivalent to reachable.

Consider a set

$$\mathcal{T}_k = \{\mathbf{x} \in \mathcal{Q}_k : f_k(\mathbf{x}, \mathbf{u}_k) \in \mathcal{E}_{k+1}, \\ \text{for some } \mathbf{u}_k \in \mathcal{U}_k\} \tag{3}$$

where

$$\mathcal{E}_{k+1} = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x} + g_k(\mathbf{w}_k)) \in \mathcal{Q}_{k+1}, \\ \forall \mathbf{w}_k \in \mathcal{W}_k\} \tag{4}$$

with boundary conditions $\mathcal{Q}_\tau = \mathcal{X}_\tau^*$ and $\mathcal{Q}_0 = \mathcal{X}_0$. Suppose $\mathcal{T}_0 \neq \emptyset$. Then $\mathbf{x}_0 \in \mathcal{T}_0 \subseteq \mathcal{X}_0$ implies $\mathbf{x}_1 \in \mathcal{Q}_1$. Similarly, if $\mathcal{T}_1 \neq \emptyset$ then $\mathbf{x}_1 \in \mathcal{T}_1 \subseteq \mathcal{Q}_1$ implies $\mathbf{x}_2 \in \mathcal{Q}_2$ etc.

Proposition 1. *The target set $\mathcal{X}_\tau^*$, $0 < \tau \in \mathbb{N}$ is reachable from all $\mathbf{x}_0 \in \mathcal{X}_0$ if and only if $\mathcal{Q}_k \subseteq \mathcal{T}_k$ for all $k \in \{0, \ldots, \tau - 1\}$.*

The sets $\mathcal{T}_k$ and $\mathcal{E}_k$ are, in principal, computable and can, in special cases, lead to the closed-form design of control functions $\mathbf{u}_k$ that ensure the target set is indeed reached.

Problem 2. *Consider a time $0 < \tau \in \mathbb{N}$ and suppose $\mathcal{X}_0$ is given and fixed. Design a control function $\mathbf{u}_k = u_k(\mathbf{x}_k) : \mathcal{X}_k \to \mathcal{U}_k$ mapping $\mathbf{x}_k$ into $\mathcal{U}_k$, for all $k \in \{0, \ldots, \tau - 1\}$ such that the following holds*

$$f_{k-1} + g_{k-1} : \mathcal{W}_{k-1} \times \{\mathcal{X}_{k-1} \times \mathcal{U}_{k-1}\} \to \mathcal{X}_k^* \tag{5}$$

for all $k \in \{1, \ldots, \tau\}$ and $\forall \mathbf{w}_k \in \mathcal{W}_k$. The sequence of sets $\{(\mathcal{X}_k^*, k)\}_{k=1}^\tau$, $0 < \tau \in \mathbb{N}$ constitutes a desired, or target, tube in $\mathbb{R}^n \times \mathbb{N}_1^\tau$ for the controlled system (1) and should be known to the control designer a priori.

We introduce the following definition.

Definition 3. *A sequence of sets $\{\mathcal{X}_k^*\}_{k=1}^\tau$, $0 < \tau \in \mathbb{N}$ is said to be **reachable given** $\mathcal{X}_0$ if there exists, at least, a single sequence of control functions $u_k : \mathcal{X}_k \to \mathcal{U}_k$, $\forall k \in \{0, \ldots, \tau - 1\}$ that solves Problem 2.*

In a different, but equivalent, notation we have it that $u_k : \mathcal{X}_k^* \to \mathcal{U}_k$ when $k \in \{1, \ldots, \tau - 1\}$.

Consider a set

$$\mathcal{T}_k^* = \{\mathbf{x} \in \mathcal{X}_k^* : f_k(\mathbf{x}, \mathbf{u}_k) \in \mathcal{E}_{k+1}^*, \\ \text{for some } \mathbf{u}_k \in \mathcal{U}_k\} \tag{6}$$

where

$$\mathcal{E}_{k+1}^* = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x} + g_k(\mathbf{w}_k)) \in \mathcal{X}_{k+1}^*, \quad \forall \mathbf{w}_k \in \mathcal{W}_k\} \tag{7}$$

where we abuse notation and let $\mathcal{X}_0^* = \mathcal{X}_0$. Suppose $\mathcal{T}_0^* \neq \emptyset$. Then $\mathbf{x}_0 \in \mathcal{T}_0^* \subseteq \mathcal{X}_0$ implies $\mathbf{x}_1 \in \mathcal{X}_1^*$. Similarly, if $\mathcal{T}_1^* \neq \emptyset$ then $\mathbf{x}_1 \in \mathcal{T}_1^* \subseteq \mathcal{X}_1^*$ implies $\mathbf{x}_2 \in \mathcal{X}_2^*$ etc.

Proposition 2. *The sequence of target sets* $\{\mathcal{X}_k^*\}_{k=1}^{\tau}$, $0 < \tau \in \mathbb{N}$ *is reachable from all* $\mathbf{x}_0 \in \mathcal{X}_0$ *if and only if* $\mathcal{T}_k^* \neq \emptyset$ *for all* $k \in \{0, \ldots, \tau - 1\}$ *and* $\mathcal{T}_0^* = \mathcal{X}_0$.

III. A Compromised System Architecture

The system introduced in the previous section is depicted in Figure 1.

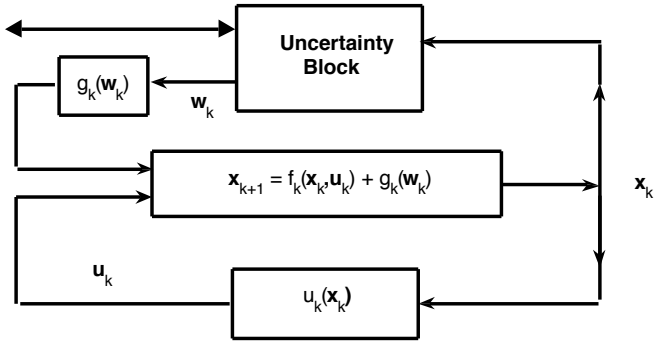


Fig. 1. A nominal robust control system.

However, suppose now an attacking, or malicious, agent places itself within the system architecture as shown in Figure 2.

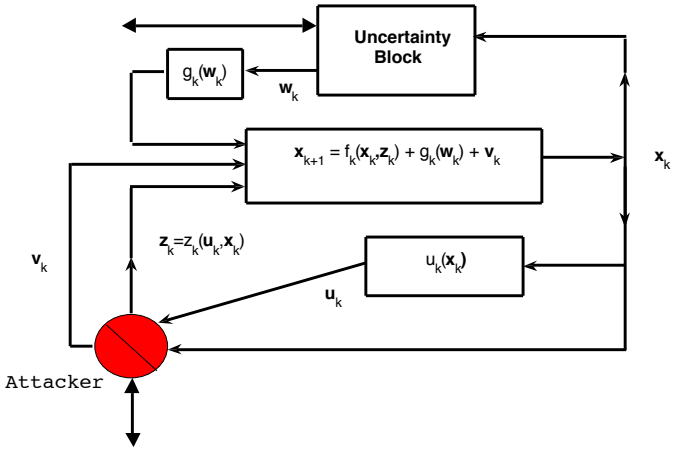


Fig. 2. A compromised robust control system. Note we will substitute $\widehat{\mathbf{x}}_k$ for $\mathbf{x}_k$ in the text of this paper in order to distinguish the trajectories of a nominal system and the compromised system for some fixed sequence of $\mathbf{u}_k$ and $\mathbf{w}_k$.

The corresponding compromised time-varying discrete-time system is then

$$\widehat{\mathbf{x}}_{k+1} = f_k(\widehat{\mathbf{x}}_k, \mathbf{z}_k) + g_k(\mathbf{w}_k) + \mathbf{v}_k \tag{8}$$

defined for $k \in \mathbb{N} \cup \{0\} = \mathbb{N}_0$ and where $\widehat{\mathbf{x}}_k \in \widehat{\mathcal{X}}_k \subseteq \widehat{\mathcal{X}} \subseteq \mathbb{R}^n$ and $\mathbf{w}_k \in \mathcal{W}_k \subseteq \mathcal{W} \subseteq \mathbb{R}^l$ are the state and disturbance vectors. Now $\mathbf{z}_k \in \mathcal{Z}_k \subseteq \mathcal{Z} \subseteq \mathcal{U} \subseteq \mathbb{R}^m$ is the *compromised control vector* and $\mathbf{v}_k \in \mathcal{V}_k \subseteq \mathcal{V} \subseteq \mathbb{R}^n$ is an *explicit attack vector*[3]. Note that we suppose $\mathcal{U}$ is the entire set of admissible control inputs to the system and thus $\mathcal{Z} \subseteq \mathcal{U}$ makes sense. Thus, the functions $f_k : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}^n$ and $g_k : \mathcal{W}_k \to \mathbb{R}^n$ are known as before and

$$f_k + g_k + \mathbf{v}_k : \mathcal{W}_k \times \{\mathbb{R}^n \times \mathcal{U}\} \times \mathcal{V}_k \to \widehat{\mathcal{X}}_{k+1} \tag{9}$$

as expected. The following definition is given for completeness.

[3]The main analysis will neglect such an input. This is not unreasonable since in practice an attacker is likely to systematically gain access to and distort only the nominal control input.

Definition 4. *Suppose $\widehat{\mathcal{X}}_0$ is given and fixed. Suppose a sequence of compromised controller functions are specified of the form $\mathbf{z}_k = z_k(\widehat{\mathbf{x}}, \mathbf{u}_k) : \widehat{\mathcal{X}}_k \times \mathcal{U}_k \to \mathcal{Z}_k$ where $\mathbf{u}_k = u_k(\widehat{\mathbf{x}}_k) : \widehat{\mathcal{X}}_k \to \mathcal{U}_k$ is a sequence of fixed nominal control vectors and where*

$$\begin{aligned}\widehat{\mathcal{X}}_{k+1} = \{\widehat{\mathbf{x}}_{k+1} \in \mathbb{R}^n : \widehat{\mathbf{x}}_{k+1} = f_k(\widehat{\mathbf{x}}_k, \mathbf{z}_k) + g_k(\mathbf{w}_k) + \mathbf{v}_k \\ \forall \widehat{\mathbf{x}}_k \in \widehat{\mathcal{X}}_k, \ \ \forall \mathbf{z}_t \in \mathcal{Z}_t, \ \forall \mathbf{u}_t \in \mathcal{U}_t, \\ \forall \mathbf{w}_t \in \mathcal{W}_t, \ \forall \mathbf{v}_t \in \mathcal{V}_t, \ \forall t \in \{0, \dots, k\}\}\end{aligned}$$

*is then called the **local set of admissible compromised state values under the sequence** $\{\mathcal{Z}_t\}_{t=0}^k$ **and** $\{\mathcal{U}_t\}_{t=0}^k$ at time k for all $k > 0$.*

If we suppose that z_k take any value in $\mathcal{U}$ at time k and suppose $\widehat{\mathcal{X}}_0$ is given and fixed then

$$\begin{aligned}\widehat{\mathfrak{X}}_{k+1} = \{\widehat{\mathbf{x}}_{k+1} \in \mathbb{R}^n : \widehat{\mathbf{x}}_{k+1} = f_k(\widehat{\mathbf{x}}_k, \mathbf{z}_k) + g_k(\mathbf{w}_k) + \mathbf{v}_k \\ \forall \widehat{\mathbf{x}}_k \in \widehat{\mathfrak{X}}_k, \ \forall \mathbf{z}_t \in \mathcal{U}, \ \forall \mathbf{w}_t \in \mathcal{W}_t, \\ \forall \mathbf{v}_t \in \mathcal{V}_t, \ \forall t \in \{0, \dots, k\}\}\end{aligned}$$

*is called the **global set of admissible compromised state values** at time k for all $k > 0$.*

The following result is given for completeness.

Proposition 3. *Suppose $\mathcal{X}_0 = \widehat{\mathcal{X}}_0$. Then $\mathfrak{X}_k \subseteq \widehat{\mathfrak{X}}_k$ and $\mathfrak{X}_k = \widehat{\mathfrak{X}}_k$ if and only if $\mathbf{v}_t = \mathbf{0}$ for all $0 \leq t \leq k-1$.*

In general, $\mathfrak{X}_k \subset \widehat{\mathfrak{X}}_k$ implies the malicious agent may be able to distort the trajectory of the system into alternative regions of the state space $\mathbb{R}^n$ than are otherwise admissible.

IV. When Can an Attack Compromise the Safety of the System?

In this section we will consider the notion of a safe operating region for a particular control system.

Definition 5. *Consider a sequence of regions $\mathcal{S}_k \subset \mathfrak{X}_k = \widehat{\mathfrak{X}}_k \subseteq \mathbb{R}^n$ for all k. We call $\mathcal{S}_k$ the **safe operating region at** k for the nominal system (1) and if $\mathbf{x}_k \notin \mathcal{S}_k$ we say the **safety of the system is compromised**.*

Note that $\cup_k \mathcal{S}_k \subset \cup_k \mathfrak{X}_k = \cup_k \widehat{\mathfrak{X}}_k \subseteq \mathbb{R}^n$. Before we look at the damage an attacker can cause under some modelling assumptions, we first look at the nominal control problem again, i.e. Problem 2.

Define a set

$$\mathcal{U}_k^+ = \{\mathbf{u} \in \mathcal{U}_k : f_k(\mathbf{x}_k, \mathbf{u}) \in \mathcal{E}_{k+1}^*, \forall \mathbf{x}_k \in \mathcal{X}_k^*\} \tag{10}$$

where we abuse notation and let $\mathcal{X}_0^* = \mathcal{X}_0$. Recall

$$\begin{aligned}\mathcal{E}_{k+1}^* = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x} + g_k(\mathbf{w}_k)) \in \mathcal{X}_{k+1}^*, \\ \forall \mathbf{w}_k \in \mathcal{W}_k\}\end{aligned} \tag{11}$$

We then have the following result.

Proposition 4. *The sequence of target sets $\{\mathcal{X}_k^*\}_{k=1}^\tau$, $0 < \tau \in \mathbb{N}$ is reachable for all $\mathbf{x}_0 \in \mathcal{X}_0$ if $\mathcal{U}_k^+ \neq \emptyset$, $\forall k$.*

The proof is omitted but note the condition $\mathcal{U}_k^+ \neq 0$ is only sufficient and not necessary to solve Problem 2. In some sense, the set $\mathcal{U}_k^+$ defined above neglects the control history $\mathbf{u}_t$, for $k \in \{0, \dots, k-1\}$. We can state a necessary and sufficient condition on a set of control inputs that ensures Problem 2 is solved.

Define the set

$$\begin{aligned}\mathcal{U}_k^* = \{\mathbf{u} \in \mathcal{U}_k : f_k(\mathbf{x}_k, \mathbf{u}) \in \mathcal{E}_{k+1}^*, \\ \forall \mathbf{x}_k \in \mathcal{F}_{k-1}(\mathbf{u}_{k-1})\}\end{aligned} \tag{12}$$

where

$$\mathcal{F}_k(\mathbf{u}_k) = f_k(\mathcal{X}_k, \mathbf{u}_k) \bigoplus g_k(\mathcal{W}_k) \tag{13}$$

where $\bigoplus$ is the set-theoretic Minkowski sum. We also let $\mathcal{F}_{-1} = \mathcal{X}_0$. Using $\mathcal{U}_k^*$ we can account for a particular history of control laws.

Proposition 5. *The sequence of target sets $\{\mathcal{X}_k^*\}_{k=1}^\tau$, $0 < \tau \in \mathbb{N}$ is reachable for all $\mathbf{x}_0 \in \mathcal{X}_0$ if and only if $\mathcal{U}_k^* \neq \emptyset$, $\forall k$.*

Obviously the following result follows from the preceding two propositions.

Corollary 1. $\mathcal{U}_k^+ \subseteq \mathcal{U}_k^*$

The set $\mathcal{U}_k^*$ is, in principal, computable and can, in special cases, lead to the closed-form design of control functions $\mathbf{u}_k$. Indeed, once $\mathcal{U}_k^*$ is computed, any $\mathbf{u}_k \in \mathcal{U}_k^*$ is sufficient. (As is typical, computation of these control input sets involves a backward recursion).

Now we examine when an attack can compromise the safety of a particular system. We make the following two assumptions.

Assumption 1. *Consider the two sets $\mathcal{N}_0^\theta$ and $\mathcal{N}_{\theta+1}^\tau$ and suppose $\mathcal{X}_0$ is given. There exists a **nominal control** $\mathbf{u}_k = u_k(\mathbf{x}_k) : \mathcal{X}_k \to \mathcal{U}_k$, $\forall k \in \{0, \dots, \tau-1\}$ such that $\mathcal{X}_k = \mathcal{X}_k^* \subseteq \mathcal{S}_k$ for all $k \in \{0, \dots, \tau\}$.*

The previous assumption states that under a nominal, i.e. non-compromised control, the system will operate within its desired target tube and within its safety region.

Assumption 2. *Consider the two sets $\mathcal{N}_0^\theta$ and $\mathcal{N}_\theta^\tau$ and suppose $\mathcal{X}_0$ is given. The attacker can hijack the nominal control inputs at times $k \in \{\theta, \dots, \tau-1\}$. The **compromised control** is $\mathbf{z}_k = z_k(\widehat{\mathbf{x}}, \mathbf{u}_k) : \widehat{\mathcal{X}}_k \times \mathcal{U}_k \to \mathcal{Z}_k$. Also, $\mathbf{v}_k \in \mathcal{V}_k = \{\mathbf{0}\}$ implying $\mathfrak{X}_k = \widehat{\mathfrak{X}}_k$ for all k and $\mathcal{X}_k = \widehat{\mathcal{X}}_k$.*

This assumption implies that an attacker can hijack the control input at some time $k = \theta$. Before this time the system is nominally controlled.

Definition 6. *Consider a sequence of regions $\mathcal{S}_k \subset \mathfrak{X}_k = \widehat{\mathfrak{X}}_k \subseteq \mathbb{R}^n$ for all k. We say the **safety of the system is compromised by an attack** if $\mathbf{x}_t \notin \mathcal{S}_t$ at some t and $\exists k \in \{0, \ldots, t-1\}$ such that $\mathbf{z}_k \neq \mathbf{u}_k$.*

Note if the safety of the system is compromised by an attack then necessarily the state of the system is outside the desired target tube since $\mathcal{X}_k^* \subseteq \mathcal{S}_k$ for all k. However, the system may be steered outside the target tube by an attacker but remain within a safe operating region for the system.

Define a set $\mathcal{A}_k = \mathfrak{X}_k \setminus \mathcal{S}_k \subset \mathfrak{X}_k = \widehat{\mathfrak{X}}_k$. This set $\mathcal{A}_k$ is a pseudo-target set for an attacker wishing to compromise the safety of the system. For simplicity, we assume it is sufficient to consider a target set $\mathcal{A}_\tau = \mathfrak{X}_\tau \setminus \mathcal{S}_\tau \subset \mathfrak{X}_\tau = \widehat{\mathfrak{X}}_\tau$.

For $k \in \{\theta, \ldots, \tau - 1\}$, consider a set

$$\mathcal{Z}_k^* = \{\mathbf{z} \in \mathcal{Z}_k : f_k(\mathbf{x}_k, \mathbf{z}) \in \mathcal{H}_{k+1}, \quad \forall \mathbf{x}_k \in \mathcal{F}_{k-1}(\mathbf{z}_{k-1})\} \tag{14}$$

where

$$\mathcal{H}_{k+1} = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x} + g_k(\mathbf{w}_k)) \in \mathcal{Q}_{k+1}, \quad \forall \mathbf{w}_k \in \mathcal{W}_k\} \tag{15}$$

with the constraint $\mathcal{Q}_\tau = \mathcal{A}_\tau$ and

$$\mathcal{F}_k(\mathbf{z}_k) = f_k(\mathcal{X}_k, \mathbf{z}_k) \bigoplus g_k(\mathcal{W}_k) \tag{16}$$

as before. We also let $\mathcal{F}_{\theta-1}(\mathbf{z}_{\theta-1}) = \mathcal{F}_{\theta-1}(\mathbf{u}_{\theta-1})$ and note that one could think of the relationship $\mathcal{Z}_k^* = \mathcal{U}_k^*$ for all $k \in \{0, \ldots, \theta - 1\}$.

Theorem 1. *The safety of the system at time $k = \tau$ can be compromised by an attacker initiating an attack at some time $k = \theta < \tau$ if and only if $\mathcal{Z}_k^* \neq \emptyset$ for all $k \in \{\theta, \ldots, \tau - 1\}$.*

The set $\mathcal{Z}_k^*$ is, in principal, computable and can, in special cases, lead to the closed-form design of attack functions $\mathbf{z}_k$. Indeed, once $\mathcal{Z}_k^*$ is computed, any $\mathbf{z}_k \in \mathcal{Z}_k^*$ is sufficient to drive the system outside its safety region so long as the necessary and sufficient condition of the preceding theorem holds. (As is typical, computation of these attack input sets involves a backward recursion).

The preceding theorem defines a necessary and sufficient set-theoretic condition on the attacker's control law given an initial and terminal attack time θ and τ respectively. Note that the time length $\tau - \theta$ obviously plays a role in whether or not the condition of the preceding theorem will hold. This preceding result leads to a straightforward definition concerning robust, secure control.

Definition 7. *If $\nexists \mathbf{x}_\theta \in \mathcal{F}_{\theta-1}(\mathbf{u}_{\theta-1})$ such that $\exists k \in \{\theta, \ldots, \tau - 1\}$ with $\mathcal{Z}_k^* = \emptyset$ then the system is said to be **securely controllable**.*

Consider a set function

$$\mathcal{P}_k(\mathcal{N}, \mathcal{L}, \mathcal{M}) = \{\mathbf{x} \in \mathcal{N} : f_k(\mathbf{x}, \mathcal{L}) \bigoplus g_k(\mathcal{W}_k) \subseteq \mathcal{M}\} \tag{17}$$

of those $\mathbf{x} \in \mathcal{N}$ such that $(f_k(\mathbf{x}, \mathbf{u}) + g_k(\mathbf{w}_k)) \in \mathcal{M}$ for every input $\mathbf{u} \in \mathcal{L}$ and $\forall \mathbf{w}_k \in \mathcal{W}_k$.

Theorem 2. *Suppose at time θ that $\mathbf{x}_\theta \in \mathcal{X}_\theta^* \subseteq \mathcal{S}_\theta$. Let $\mathcal{F}_{\theta-1}(\mathcal{Z}_{\theta-1}) = \mathcal{X}_\theta^*$ for notational simplicity. If*

$$\mathcal{P}_t(\mathcal{F}_{t-1}(\mathcal{Z}_{t-1}) \cap \mathcal{S}_t, \mathcal{Z}_t, \mathcal{S}_{t+1}) \neq \mathcal{F}_{t-1}(\mathcal{Z}_{t-1}) \cap \mathcal{S}_t \tag{18}$$

for some $t \geq \theta$ then the system is securely controllable so long as $\tau \leq t$.

The preceding theorem is the main result and provides a condition for a system designer which results in a necessary and sufficient lower bound on the number of consecutive attack signals required to compromise the safety of the system.

Corollary 2. *Suppose at time θ that $\mathbf{x}_\theta \in \mathcal{X}_\theta^* \subseteq \mathcal{S}_\theta$. Let $\mathcal{F}_{\theta-1}(\mathcal{Z}_{\theta-1}) = \mathcal{X}_\theta^*$ for notational simplicity. Consider the optimization problem*

$$\begin{aligned} \tau^* = \underset{\tau}{\operatorname{argmax}} \ & \tau \\ \text{such that } & \mu\left(\bigcap_{t=\theta}^{\tau} (\mathcal{F}_{t-1}(\mathcal{Z}_{t-1}) \cap \mathcal{S}_t)\right) \\ & -\mu\left(\bigcap_{t=\theta}^{\tau} \mathcal{P}_t(\mathcal{F}_{t-1}(\mathcal{Z}_{t-1}) \cap \mathcal{S}_t, \mathcal{Z}_t, \mathcal{S}_{t+1}\right) > 0 \end{aligned} \tag{19}$$

where the operation $\mu(\cdot)$ gives the Lebesque measure of its argument. Then there exists a suitable attack sequence of length $\tau^* - \theta$ that can compromise the safety of the system at time τ^*. Conversely, there is no attack sequence of a shorter duration that can compromise the safety of the system.

The preceding two results, for example, provide computable functions on the known system parameters, e.g. the model and the input constraints etc, which can aid the system designer in securing the control system. In practice, if an attacker wanted to drive the state outside some safety region and the attacker had an arbitrarily long time period within which to distort the system controller then (apart from some special cases) it is unreasonable to expect this to be impossible.

The preceding two results lead to a necessary and sufficient lower bound on the time period in which an attacker can take the state from a safety region to a unsafe region. The system designer can then design an attack detection algorithm or a monitoring process etc that attempts to verify the integrity of the controller within this time period. It has also been discussed in, e.g., [8] that an attacker may only intermittently be able to distort a control signal. Thus, a designer may be able to configure the system such that the lower bound introduced in this paper is increased.

V. Concluding Remarks

This paper briefly explored the classical notion of control system reachability as applied within a secure control theoretic framework.

Note that all the set-theoretical results introduced in this paper are, in principal, computable; e.g. it is possible to design controller and attack input sets that drive the state of the system to the respective target sets (so long as those sets are reachable). For linear time-varying systems, where the control, attack and disturbance sets are ellipsoidal etc, the set-theoretical results introduced in this work can be well-approximated using ellipsoidal calculus.

Acknowledgement

This work is supported by NICTA which is funded by the Australian Government via the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. This work is also supported by the US Air Force via the Asian Office of Aerospace Research and Development (AOARD) and contract USAF-AOARD-10-4102.

References

[1] G.N. Nair, F. Fagnani, S. Zampieri, and R.J. Evans. Feedback control under data rate constraints: an overview. *Proceedings of the IEEE*, 95(1):108–137, March 2007.

[2] J.P. Hespanha, P. Naghshtabrizi, and Y. Xu. A survey of recent results in networked control systems. *Proceedings of the IEEE*, 95(1):138–162, March 2007.

[3] A.S. Matveev and A.V. Savkin. *Estimation and Control over Communication Networks*. Birkhauser, Boston, 2008.

[4] A.A. Cárdenas, S. Amin, and S.S. Sastry. Secure control: Towards survivable cyber-physical systems. In *Proceedings of the 1st International Workshop on Cyber-Physical Systems*, pages 495–500, Beijing, China, June 2008.

[5] K. O'Brien. A silent attack, but not a subtle one. *The New York Times*, September 26 2010.

[6] J. Meserve. Staged cyber attack reveals vulnerability in power grid. *CNN*, September 26 2007.

[7] S. Gorman. Electricity grid in U.S. penetrated by spies. *The Wall Street Journal*, April 8 2009.

[8] S. Amin, A.A. Cárdenas, and S.S. Sastry. Safe and secure networked control systems under denial-of-service attack. In *Hybrid Systems: Computation and Control*, pages 31–45. Lecture Notes in Computer Science, Springer, April 2009.

[9] P. Mohajerin, M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson. Cyber attack in a two-area power system: Impact identification using reachability. In *Proc. of the 2010 American Control Conference*, Baltimore, MD, June 2010.

[10] P. Mohajerin, M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson. A robust policy for automatic generation control cyber attack in two area power network. In *Proc. of the 49th IEEE Conf. on Decision and Control*, pages 5973–5978, Atlanta, GA, December 2010.

[11] Y. Mo and B. Sinopoli. False data injection attacks in cyber physical systems. In *Proceedings of the 1st Workshop on Secure Control Systems*, Stockholm, Sweden, April 2010.

[12] S. Sundaram and C.N. Hadjicostis. Distributed function calculation via linear iterations in the presence of malicious agents - part I: Attacking the network. In *Proceedings of the 2008 American Control Conference*, pages 1350–1355, Seatle, WA, USA, June 2008.

[13] S. Sundaram and C.N. Hadjicostis. Distributed function calculation via linear iterations in the presence of malicious agents - part II: Overcoming malicious behavior. In *Proceedings of the 2008 American Control Conference*, pages 1356–1361, Seatle, WA, USA, June 2008.

[14] Y. Liu, M.K. Reiter, and P. Ning. False data injection attacks against state estimation in electric power grids. In *Proceedings of the 16th ACM Conference on Computer and Communications Security*, pages 21–32, November.

[15] R.B. Bobba, K.M. Rogers, Q. Wang, H. Khurana, K. Nahrstedt, and T.J. Overbye. Detecting false data injection attacks on DC state estimation. In *Proceedings of the 1st Workshop on Secure Control Systems*, Stockholm, Sweden, April 2010.

[16] H. Sandberg, A. Teixeira, and K.H. Johansson. On security indices for state estimators in power networks. In *Proceedings of the 1st Workshop on Secure Control Systems*, Stockholm, Sweden, April 2010.

[17] O. Kosut, L. Jia, R.J. Thomas, and L. Tong. Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures. In *Proc. of the 1st IEEE Int. Conference on Smart Grid Communications*, pages 220–225, Gaithersburg, MD, USA, October 2010.

[18] G. Dán and H. Sandberg. Stealth attacks and protection schemes for state estimators in power systems. In *Proceedings of the 1st IEEE International Conference on Smart Grid Communications*, pages 214–219, Gaithersburg, MD, USA, October 2010.

[19] B. Zhu and S. Sastry. SCADA-specific intrusion detection/prevention systems: A survey and taxonomy. In *Proceedings of the 1st Workshop on Secure Control Systems*, Stockholm, Sweden, April 2010.

[20] Le Xie, Yilin Mo, and B. Sinopoli. False data injection attacks in electricity markets. In *Proceedings of the 1st IEEE International Conference on Smart Grid Communications*, pages 226–231, Gaithersburg, MD, USA, October 2010.

[21] A. Teixeira, H. Sandberg, and K.H. Johansson. Networked control systems under cyber attacks with applications to power networks. In *Proc. of the 2010 American Control Conference*, pages 3690–3696, Baltimore, MD, June 2010.

[22] F. Pasqualetti, R. Carli, A. Bicchi, and F. Bullo. Identifying cyber attacks via local model information. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 5961–5966, Atlanta, GA, USA, December 2010.

[23] S. Sundaram, M. Pajic, C.N. Hadjicostis, R. Mangharam, and G.J. Pappas. The wireless control network: Monitoring for malicious behavior. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 5979–5984, Atlanta, GA, USA, December 2010.

[24] A. Teixeira, S. Amin, H. Sandberg, K.H. Johansson, and S.S. Sastry. Cyber security analysis of state estimators in electric power systems. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 5991–5998, Atlanta, GA, USA, December 2010.

[25] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli. False data injection attacks against state estimation in wireless sensor networks. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 5967–5972, Atlanta, GA, USA, December 2010.

[26] S. Zheng, T. Jiang, and J.S. Baras. Robust state estimation under false data injection in distributed sensor networks. In *Proceedings of the 2010 IEEE Global Telecommunications Conference*, Miami, FL, USA, December 2010.

[27] A.N. Bishop and A.V. Savkin. On false-data attacks in robust multi-sensor-based estimation. In *Proceedings of the 9th IEEE International Conference on Control and Automation*, Santiago, Chile, December 2011.

[28] R. Isermann. *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. Springer Verlag, 2006.

[29] H. Witsenhausen. A minimax control problem for sampled linear systems. *IEEE Transactions on Automatic Control*, 13(1):5–21, February 1968.

[30] D.P. Bertsekas and I.B. Rhodes. On the minimax reachability of target sets and target tubes. *Automatica*, 7(2):233–247, March 1971.

[31] J. Glover and F. Schweppe. Control of linear dynamic systems with set constrained disturbances. *IEEE Transactions on Automatic Control*, 16(5):411–423, October 1971.

Remarks on the Cramer-Rao Inequality for Doppler-Based Target Parameter Estimation

Adrian N. Bishop [#1] and Matthew Smith [*2]

[#] *National ICT Australia (NICTA) and the Australian National University (ANU)*
[1] adrian.bishop@ieee.org
[*] *CEA Technologies*
[2] matthew.smith@cea.com.au

***Abstract*—This paper outlines the problem of multi-static Doppler-based target position and velocity estimation. The Fisher information matrix is derived given a separate target illuminator and then given a target-based isotropic signal emission. Some remarks concerning the Cramer-Rao inequality and its relationship to the estimation problem are given. Some results concerning the placement of the receivers are given and some open problems are discussed.**

I. Introduction

Using Doppler-shifts for position and velocity estimation has a long history; see e.g. [1]–[4]. Consider a scenario with n_s sensors, and n_o transmitters; see Figure 1. The target has an unknown position and velocity $\mathbf{x} = [\mathbf{p}^\top \ \mathbf{v}^\top]^\top \in \mathbb{R}^4$. The position of the sensors is given by $\mathbf{s}_i = [s_{i,1} \ s_{i,2}]^{\mathrm{T}} \in \mathbb{R}^2, \forall i \in \{1, \ldots, n_s\}$ while the position of the transmitters is given by $\mathbf{o} = [o_1 \ o_2]^{\mathrm{T}} \in \mathbb{R}^2$. We let $\phi_i \in [0, 2\pi)$ denote the bearing to the target at the i^{th} sensor and $\theta_i \in [0, 2\pi)$ denote the bearing to the target at the transmitter. The angle subtended at the target by two sensors i and j is denoted by $\vartheta_{ij} = \vartheta_{ji} \in [0, \pi]$ similarly the angle subtended at the target by two transmitters i and j is denoted by $\varphi_{ij} = \varphi_{ji} \in [0, \pi]$.

The measured Doppler-shift is $\hat{f}_{ij}$ at the i^{th} sensor and is caused by a target reflection due to a signal generated by the j^{th} transmitter. This frequency shift can be approximated by

$$\hat{f}_{ij} = f_{ij} + w_{ij} \tag{1a}$$

$$= \frac{f_{c,ij}}{2c}\left(\frac{(\mathbf{p}-\mathbf{s}_i)^T}{\|\mathbf{p}-\mathbf{s}_i\|} + \frac{(\mathbf{p}-\mathbf{o}_j)^T}{\|\mathbf{p}-\mathbf{o}_j\|}\right)\mathbf{v} + w_{ij} \tag{1b}$$

$$= \zeta_{ij}(\mathbf{p})\mathbf{v} + w_{ij} \tag{1c}$$

where c is the speed of light (or signal propagation) and $\|\cdot\|$ is the standard Euclidean vector norm and $f_{c,ij}$ is the carrier frequency of employed by this transmitter sensor pair. Finally, w_{ij} is a zero mean Gaussian random variable with known variance σ_{ij}^2. For notation simplicity, we employ the the two dimensional index $\lambda \triangleq (j, i)$ and following normalization $\frac{f_{c,\lambda}}{c} = \frac{f_{c,ij}}{c} \triangleq 1$. The set of all transmitter/sensor pairs is denoted $\Lambda \subseteq [1, n_s] \times [1, n_o]$ and we employ a lexical order; $(\lambda(1), \lambda(2), \ldots, \lambda(n)) \equiv \Lambda$, where n is the total number of transmitter/sensor pairs.

In the following we delineate four different, but related, RADAR/sensor network configurations. Our focus is on the types of measurements generated by the sensor and the time-energy characteristic of each configuration. The speed of signal propagation in all scenarios is a known constant c.

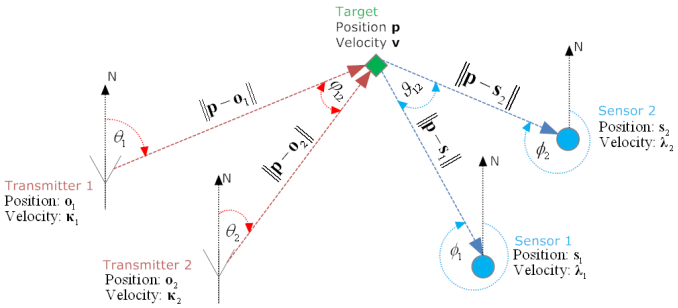


Fig. 1. A typical multi-static target state estimation scenario with one target, two sensors and two transmitters.

1) **Passive**: Here the target itself transmits a continuous-wave signal at some frequency f_c which is received by n sensors in the environment. Sensor measurements are some combination of target bearing and Doppler frequency shift. Bearing measurements require a hardware array of antennas. The advantage of Doppler-only-based sensors lies in the simplicity of the sensors since no hardware array is needed to determine the frequency shift. The obvious advantage of passive tracking systems is that a passive sensor does not indicate its location to any electronic listening devices. However, in contrast to the *active* sensor network techniques described later, there are numerous disadvantages to passive sensor networks:
   a) Unless the target is friendly, knowledge of the carrier frequency f_c is not immediate and must be estimated.
   b) High value and/or intelligent targets will only transmit *occasionally* - if at all, and may *vary* the transmitted waveform (and hence f_c) on each occasion.
   c) The energy received by the sensor network is dictated by the non-cooperating target.
2) **Mono-static**: Here each transmitter illuminates the *entire* area of interest and each transmitter is paired with a *particular* sensor. Each transmitted signal is reflected by the target and received by the prescribed sensor. The configuration is called mono-static whenever each

transmitter/sensor pair is collocated. The advantage of all active sensor networks (mono-static or otherwise) is that - in addition to angular and Doppler information, target range information is available to the sensors. In contrast to passive sensor networks, the transmitter frequency f_c is assumed to be known to the receiving sensor. Moreover, the transmitted signal and illumination pattern can be sophisticated in its design. These system design parameters are at the discretion of the sensor network.

3) **Bi-static**: A bi-static sensor network is similar to a mono-static network except that each of the transmitter/sensor pairs are not collocated. Bi-static networks can exploit network geometries by deploying the sensors "near" the target, thus improving the received signal-to-noise ratio. This is turn improves the accuracy and resolution capabilities of the network. Furthermore, in an electronic warfare context, a bi-static network does not give away the location of any of its sensors to electronic listening devices. The transmitted signal can encode additional information such as the transmitter's location, etc. Alternatively, the illuminator may be a so-called transmitter-of-opportunity such as a television broadcaster etc.
4) **Multi-static**: Each transmitter $j \in [1, n_o]$ illuminates some portion of the area of interest, which is then reflected by the target and received at a number $n_j \in [1, n_s]$ of sensors. By employing modern Digital Array RADAR (DAR) techniques, two or more transmitters and/or two or more receivers may in fact be collocated. A bi-static network is obviously a special case of multi-static network.

In summary we note the following:

1) Active transmission imposes considerable energy requirements. This is particularly troublesome in a mobile sensor network where the weight and energy consumption of each sensor needs to be managed efficiently.
2) By employing sensors that measure Doppler-shift-only, the complexity in the the receivers is typically much less than those sensors that measure bearing and/or range.
3) Target Doppler velocity measurements are easy to generate in an active network.
4) Bi-static and multi-static sensor networks can exploit geometries to improve signal-to-noise ratio characteristics.
5) In bi-static and multi-static sensor networks the receivers are passive sensors in this scenario and have reduced energy requirements in comparison to the illuminator. In these networks the system designer can seek the advantages of both active and passive radar systems while minimizing the disadvantages of both individually.

II. Problem Scenario

As noted in the previous section we consider broadly three different network configurations.

A. Multi-Static Scenario

From (1b) we have

$$\hat{f}_\lambda = \tfrac{1}{2}\big(\mathbf{u}_{\lambda_1} + \mathbf{u}_{\lambda_2}\big)^\top \mathbf{v} + w_{ij} \tag{2a}$$

where

$$\mathbf{u}_{\lambda_1} := \frac{\mathbf{p} - \mathbf{s}_i}{\|\mathbf{p} - \mathbf{s}_i\|} = \begin{bmatrix} \cos(\phi_{\lambda_1}) \\ \sin(\phi_{\lambda_1}) \end{bmatrix} \tag{2b}$$

$$\mathbf{u}_{\lambda_2} := \frac{\mathbf{p} - \mathbf{o}_j}{\|\mathbf{p} - \mathbf{o}_j\|} = \begin{bmatrix} \cos(\theta_{\lambda_2}) \\ \sin(\theta_{\lambda_2}) \end{bmatrix} \tag{2c}$$

are the unit vectors directed towards the target from the sensor and transmitter respectively. Stacking the measurements from all transmitter/sensor pairs we obtain

$$\hat{\mathbf{f}} = \mathbf{f} + \mathbf{w} \tag{3}$$

$$= \big[f_{\lambda(1)}, f_{\lambda(2)}, \ldots, f_{\lambda(n)}\big]^T + \big[w_{\lambda(1)}, w_{\lambda(2)}, \ldots, w_{\lambda(n)}\big]^T \tag{4}$$

which obeys $\hat{\mathbf{f}} \sim \mathcal{N}(\mathbf{f}, \mathbf{R}_f)$ with covariance matrix $\mathbf{R}_f = \mathrm{diag}(\sigma^2_{\lambda(1)}, \ldots, \sigma^2_{\lambda(n)})$ Note that

$$\hat{\mathbf{f}} = \mathrm{diag}(\zeta_{\lambda(1)}(\mathbf{p}), \ldots, \zeta_{\lambda(n)}(\mathbf{p}))\mathbf{v} + \mathbf{w} \tag{5a}$$

$$= \mathbf{H}(\mathbf{p})\mathbf{v} + \mathbf{w} \tag{5b}$$

which highlights the fact $\hat{\mathbf{f}}$ is a vector of nonlinear measurements in $\mathbf{p}$ but is linear in the target velocity $\mathbf{v}$.

Note that the illuminators in this scenario can be quite sophisticated in their design of the signal waveform and illumination pattern. For example, the illumination signal can encode the position of the transmitter along with other relevant signal information. Alternatively, the illuminator may be a so-called transmitter-of-opportunity such as a television broadcaster or cellular phone tower. The receivers are passive sensors in this scenario and can be relatively simple devices in comparison to the illuminator. This scenario arguably seeks the advantages of both active and passive radar systems while minimizing the disadvantages of both individually.

We note some practical problems that arise in multi-static scenarios owing to the distributed sensing nature of multi-static configurations. Firstly, difficulties with incoherence when each sensor is running a local oscillator independently of the other sensors and even the illuminator may exist. Furthermore, signal measurements may arrive at different sensors asynchronously. In addition, the sensors must communicate either to a central estimator or amongst themselves in order to make use of the received measurements. Asynchronous polling, communication bandwidths, energy consumption etc can all contribute to significant (often underestimated) communication problems with multi-static networks in practice.

B. Mono-Static Scenario

In a mono-static configuration we have an equal number of sensors and transmitters $n_s = n_o := n$ and each

sensor/transmitter pair is collocated. The following equality conditions are immediate: for each $\lambda \in \Lambda$

$$\mathbf{u}_{\lambda_1} = \mathbf{u}_{\lambda_2} \quad (=: \mathbf{u}_\lambda) \tag{6a}$$
$$\phi_{\lambda_1} = \theta_{\lambda_2} \tag{6b}$$
$$\hat{\mathbf{f}}_\lambda = \mathbf{u}_\lambda^T \mathbf{v} + \mathbf{w} \tag{6c}$$

C. Passive Configuration

In this configuration we have $n = n_s$ and the target itself is the (isotropic) emitter. Observe that the measurement equation (2) for the passive configuration coincides with the mono-static case (6c). Geometrically speaking this should be no suprise as each passive sensor measures only that portion of the isotropic emission that is directed towards the target.

III. The Cramer-Rao Inequality

If $\mathcal{I}(\mathbf{x})$ is the Fisher information matrix then the Cramer-Rao inequality lower bounds the variance achievable by an unbiased estimator. For an *unbiased* estimate $\widehat{\mathbf{x}}$ of $\mathbf{x}$ we find

$$\mathrm{E}\left[(\widehat{\mathbf{x}} - \mathbf{x})(\widehat{\mathbf{x}} - \mathbf{x})^\top\right] \geq \mathcal{I}(\mathbf{x})^{-1} \tag{7}$$

If $\mathcal{I}(\mathbf{x})$ is singular then (in general) no unbiased estimator for $\mathbf{x}$ exists with a finite variance [5]. If (7) holds with equality, for some unbiased estimate $\widehat{\mathbf{x}}$, then the estimator is called *efficient* and the parameter estimate $\widehat{\mathbf{x}}$ is unique [5]. However, even if $\mathcal{I}(\mathbf{x})$ is non-singular then it is not practically guaranteed that an unbiased estimator can be recognized. Alternatively, if an unbiased estimator can be realized, it is not guaranteed that an efficient estimator exists [6]. The design of specific estimation algorithms is not the immediate goal of this work. However, it is obvious that an unbiased estimator $g(\widehat{\mathbf{f}})$ of $\mathbf{x}$ is one which obeys

$$\mathrm{E}[g(\widehat{\mathbf{f}})] = \mathbf{x} = \frac{1}{(2\pi)^{n/2}|\mathbf{R_f}|^{1/2}} \int_{\mathbb{R}^n} g(\widehat{\mathbf{f}}) e^{-\frac{1}{2}\mathbf{w}^\top \mathbf{R_f}^{-1}\mathbf{w}} \, \mathrm{d}\mathbf{w} \tag{8}$$

where again $\widehat{\mathbf{f}} = \mathbf{f} + \mathbf{w}$. Starting with this constraint, there are a number of strategies of designing unbiased estimators; see [6]. We will not explore this concept further. The condition (7) says nothing about the performance and realizability of biased estimators. That is, in order to use (7) we must consider only unbiased estimators [5]. The $(i,j)^{th}$ element of $\mathcal{I}(\mathbf{x})$ is given by

$$\mathcal{I}_{i,j}(\mathbf{x}) = \mathrm{E}\left[\frac{\partial}{\partial x_i} \ln\left(f_{\widehat{\mathbf{p}}}(\widehat{\mathbf{p}};\mathbf{x})\right) \frac{\partial}{\partial x_j} \ln\left(f_{\widehat{\mathbf{p}}}(\widehat{\mathbf{p}};\mathbf{x})\right)\right] \tag{9}$$

where $[\mathbf{p}^\top \ \mathbf{v}^\top]^\top = [x_1 \ \ldots \ x_4] \in \mathbb{R}^4$ and $f_{\mathbf{f}}(\mathbf{f};\mathbf{x})$ is the Gaussian likelihood function. We then easily find $\mathcal{I}(\mathbf{x}) = \nabla_{\mathbf{x}}\mathbf{f}(\mathbf{x})^{\mathrm{T}}\mathbf{R}_{\mathbf{f}}^{-1}\nabla_{\mathbf{x}}\mathbf{f}(\mathbf{x})$. The Fisher information metric characterizes the nature of the likelihood function. If the likelihood function is sharply peaked then the true parameter value is easier to estimate from the measurements than if the likelihood function is flatter.

A. General Results Concerning Each Scenario

The Fisher information matrix is given by (11). We use $\mathcal{I}(\mathbf{x})$, $\mathcal{I}(\mathbf{p})$ and $\mathcal{I}(\mathbf{v})$ to denote the Fisher information matrix defined by considering only the parameters $\mathbf{x}$, $\mathbf{p}$ and $\mathbf{v}$ respectively. Both $\mathcal{I}(\mathbf{p})$ and $\mathcal{I}(\mathbf{v})$ turn out to be principal submatrices of $\mathcal{I}(\mathbf{x})$. In all cases, independent measurements from additional sensors in general positions will never decrease the total information in each $\mathbf{x}$, $\mathbf{p}$ and $\mathbf{v}$.

Proposition 1. *The condition $n \geq 4$ is a necessary condition for $\mathcal{I}(\mathbf{x})$ to be non-singular.*

Proof: Recall that $\mathcal{I}(\mathbf{x}) = \nabla_{\mathbf{x}}\mathbf{f}(\mathbf{x})^{\mathrm{T}}\mathbf{R}_{\mathbf{f}}^{-1}\nabla_{\mathbf{x}}\mathbf{f}(\mathbf{x})$ or given $\mathbf{R_f} = \mathrm{diag}\left(\sigma_1^2, \ldots, \sigma_n^2\right)$ we have

$$\mathcal{I}(\mathbf{x}) = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \nabla_{\mathbf{x}}\mathbf{f}_i^\top \cdot \nabla_{\mathbf{x}}\mathbf{f}_i \tag{12}$$

which is a sum of matrices each with rank at most 1. Now a well-known result states that a rank-k matrix can be written as the sum of k rank-1 matrices but not fewer. This immediately implies our result and completes the proof. ∎

Proposition 2. *The condition $n \geq 2$ is a necessary condition for $\mathcal{I}(\mathbf{p})$ to be non-singular.*

Proposition 3. *The condition $n \geq 2$ is a necessary condition for $\mathcal{I}(\mathbf{v})$ to be non-singular.*

Proposition 4. *The following statements concerning efficient estimators hold.*

1) *If n is finite then no efficient estimator exists for $\mathbf{x}$.*
2) *If $\mathbf{p}$ is known and n is finite then an efficient estimator for $\mathbf{v}$ exists and is given by the standard linear maximum likelihood estimator.*
3) *If $\mathbf{v}$ is known and n is finite then no efficient estimator exists for $\mathbf{p}$.*

Proof: This result follows from a general result concerning efficient estimators given in [6]. ∎

We do not consider the design of unbiased (but inefficient) estimators for either $\mathbf{p}$ or $\mathbf{x}$ in this work. However, we know that no unbiased estimator for $\mathbf{x}$ exists with a finite variance when $n < 4$. Similarly, no unbiased estimator for $\mathbf{p}$ or $\mathbf{v}$ exists with a finite variance when $n < 2$.

B. Discussion on the Cramer-Rao Bound

The Cramer-Rao inequality assumes an unbiased estimation algorithm and an estimator which achieves the inequality is called an efficient estimator. An efficient estimator does not exist for $\mathbf{x}$ or $\mathbf{p}$ when n is finite but does exist for $\mathbf{v}$ when $\mathbf{p}$ is given. Even if an efficient estimator does not exist then it may be possible to design an unbiased estimator. This possibility is not explored in this work. In practice, a system designer may be constrained in their choice of parameter estimator. Likely, the estimation technique used in practice will be biased [7],

$$\mathcal{I}(\mathbf{x}) = \begin{bmatrix} \mathcal{I}(\mathbf{p}) & \cdot \\ \cdot & \mathcal{I}(\mathbf{v}) \end{bmatrix} = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \begin{bmatrix} \mathcal{I}_i(\mathbf{p}) & \cdot \\ \cdot & \mathcal{I}_i(\mathbf{v}) \end{bmatrix} \tag{10}$$

$$= \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \begin{bmatrix} \left(\frac{\partial}{\partial x_1} f_i\right)^2 & \frac{\partial}{\partial x_1} f_i \cdot \frac{\partial}{\partial x_2} f_i & \frac{\partial}{\partial x_1} f_i \cdot \frac{\partial}{\partial x_3} f_i & \frac{\partial}{\partial x_1} f_i \cdot \frac{\partial}{\partial x_4} f_i \\ \frac{\partial}{\partial x_1} f_i \cdot \frac{\partial}{\partial x_2} f_i & \left(\frac{\partial}{\partial x_2} f_i\right)^2 & \frac{\partial}{\partial x_2} f_i \cdot \frac{\partial}{\partial x_3} f_i & \frac{\partial}{\partial x_2} f_i \cdot \frac{\partial}{\partial x_4} f_i \\ \frac{\partial}{\partial x_3} f_i \cdot \frac{\partial}{\partial x_1} f_i & \frac{\partial}{\partial x_3} f_i \cdot \frac{\partial}{\partial x_2} f_i & \left(\frac{\partial}{\partial x_3} f_i\right)^2 & \frac{\partial}{\partial x_3} f_i \cdot \frac{\partial}{\partial x_4} f_i \\ \frac{\partial}{\partial x_4} f_i \cdot \frac{\partial}{\partial x_1} f_i & \frac{\partial}{\partial x_4} f_i \cdot \frac{\partial}{\partial x_2} f_i & \frac{\partial}{\partial x_4} f_i \cdot \frac{\partial}{\partial x_3} f_i & \left(\frac{\partial}{\partial x_4} f_i\right)^2 \end{bmatrix} \tag{11}$$

[8]. For example, even the well-known maximum likelihood localization techniques are only asymptotically unbiased and efficient, i.e. require the number of sensors to approach infinity. The Cramer-Rao bound for unbiased estimators is still an interesting benchmark with which intuitively pleasing results and performance measures can be derived. However, these results can only be considered as a guide.

It is interesting that the variance (or mean-square-error) of an estimate can sometimes be made smaller at the expense of increasing the bias [9]. The work of [10], [11] explores the concept of bias-variance trade offs in estimation. In [5], [11] a *biased* Cramer-Rao inequality and in [10], [11] a *uniform* Cramer-Rao inequality are developed and can be used to study this so-called bias-variance trade off. These ideas are yet to be fully explored in the localization and target tracking literature.

IV. On the Fisher Information for Velocity Estimation

Doppler-based measurements are often used to estimate the target velocity alone. In this section we explore the relationship between the transmitter and sensor positions and the velocity estimation error lower bound defined by $\mathcal{I}_\lambda(\mathbf{v})$. To this end, consider the sub-matrix

$$\mathcal{I}(\mathbf{v}) = \sum_{\lambda} \frac{\rho_\lambda^2}{\sigma_\lambda^2} \begin{bmatrix} \cos^2\left(\frac{\phi_{\lambda_1}+\theta_{\lambda_2}}{2}\right) & \frac{\sin(\phi_{\lambda_1}+\theta_{\lambda_2})}{2} \\ \frac{\sin(\phi_{\lambda_1}+\theta_{\lambda_2})}{2} & \sin^2\left(\frac{\phi_{\lambda_1}+\theta_{\lambda_2}}{2}\right) \end{bmatrix} \tag{13}$$

where $\rho_\lambda^2 := \cos^2\left(\frac{\phi_{\lambda_1}-\theta_{\lambda_2}}{2}\right)$ and $f_c/c \triangleq 1$. This simplifies to

$$\mathcal{I}(\mathbf{v}) = \frac{1}{\sigma^2} \sum_{i=1}^{n} \begin{bmatrix} \cos^2(\phi_i) & \frac{\sin(2\phi_i)}{2} \\ \frac{\sin(2\phi_i)}{2} & \sin^2(\phi_i) \end{bmatrix} \tag{14}$$

in the mono-static or passive scenario with $\sigma_i = \sigma_j = \sigma$ for all i, j and $f_c/c \triangleq 1$.

Theorem 1. *The determinant of the Fisher information matrix $\mathcal{I}_\lambda(\mathbf{v})$ in (13) is given by*

$$\det\big(\mathcal{I}(\mathbf{v})\big) = \frac{1}{4}\left[\Big(\sum_{\lambda\in\Lambda} \frac{\rho_\lambda^2}{\sigma_\lambda^2}\Big)^2 - \Big(\sum_{\lambda\in\Lambda} \frac{\rho_\lambda^2}{\sigma_\lambda^2}\cos(\phi_{\lambda_1}+\theta_{\lambda_2})\Big)^2 - \Big(\sum_{\lambda\in\Lambda} \frac{\rho_\lambda^2}{\sigma_\lambda^2}\sin\big(\phi_{\lambda_1}+\theta_{\lambda_2}\big)\Big)^2\right] \tag{15}$$

where $\rho_\lambda^2 := \cos^2\left(\frac{\phi_{\lambda_1}-\theta_{\lambda_2}}{2}\right)$.

Proof: Firstly we write,

$$\mathcal{I}(\mathbf{v}) = \tfrac{1}{2}\begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{16}$$

$$a = \sum_{\lambda} \frac{\rho_\lambda^2}{\sigma_\lambda^2}\big(1+\cos(\phi_{\lambda_1}+\theta_{\lambda_2})\big) \tag{17}$$

$$b = \sum_{\lambda} \frac{\rho_\lambda^2}{\sigma_\lambda^2}\sin(\phi_{\lambda_1}+\theta_{\lambda_2}) \tag{18}$$

$$c = \sum_{\lambda} \frac{\rho_\lambda^2}{\sigma_\lambda^2}\sin(\phi_{\lambda_1}+\theta_{\lambda_2}) \tag{19}$$

$$d = \sum_{\lambda} \frac{\rho_\lambda^2}{\sigma_\lambda^2}\big(1-\cos(\phi_{\lambda_1}+\theta_{\lambda_2})\big) \tag{20}$$

Applying the determinant formula $\det\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = ad - bc$ we obtain two terms; the first term "ad" may be reduced by completing the square:

$$\text{First Term} = \Big(\sum_{\lambda} \frac{\rho_\lambda^2}{\sigma_\lambda^2}\Big)^2 - \Big(\sum_{\lambda} \frac{\rho_\lambda^2}{\sigma_\lambda^2}\cos(\phi_{\lambda_1}+\theta_{\lambda_2})\Big)^2 \tag{21}$$

Observe that the above is the first two terms in (15). The third term in (15) is immediate from the second term "bc" in the determinate formula. ∎

In the mono-static, or passive scenario, with $\sigma_i = \sigma_j = \sigma$ the following corollary is useful.

Corollary 1. *Let ϕ_i, $\forall i \in \{1, \ldots, n\}$ denote the angular positions of the sensors. The following are equivalent expressions for the Fisher information determinant* $\det(\mathcal{I}(\mathbf{v}))$

$$\text{(i).}\quad \det(\mathcal{I}(\mathbf{v})) = \frac{1}{4\sigma^4} \times \left[n^2 - \left(\sum_{i=1}^{n}\cos(2\phi_i)\right)^2 - \left(\sum_{i=1}^{n}\sin(2\phi_i)\right)^2\right] \tag{22}$$

$$\text{(ii).}\quad \det(\mathcal{I}(\mathbf{v})) = \frac{1}{\sigma^4}\sum_{\mathcal{S}}\sin^2(\phi_j - \phi_i), \quad j > i \tag{23}$$

where $\mathcal{S} = \{\{i, j\}\}$ is defined as the set of all combinations of i and j with $i, j \in \{1, \ldots, n\}$ and $j > i$.

Proof: Part (ii) follows from [12]. ∎

In both the multi-static and mono-static scenarios we state that angular sensor positions which maximize the determinant

(15) or (22) respectively generate an optimal sensor-target geometry for velocity estimation using Doppler.

Theorem 2. *For the multi-static sensor configuration with* $\sigma_\lambda^2 \equiv \sigma^2$ *the Fisher information determinant* (15) *is upper bounded by* $\frac{n^2}{4\sigma^2}$. *This bound is achieved if and only if the angular positions of the sensors and transmitters are such that*

$$\forall \lambda \in \Lambda : \ \phi_{\lambda_1} = \theta_{\lambda_2} mod\ 2\pi \tag{24}$$

and

$$\sum_{\lambda\in\Lambda} \cos(2\phi_\lambda) = 0 \quad and \quad \sum_{\lambda\in\Lambda} \sin(2\phi_\lambda) = 0 \tag{25}$$

holds.

Theorem 3. *For the passive sensor configuration with* $\sigma_\lambda^2 \equiv \sigma^2$, *the Fisher Information determinant* (22) *is upper bounded by* $\frac{n^2}{4\sigma^2}$. *This upper bound is achieved if and only if the angular positions of the sensors are such that the equations*

$$\sum_{\lambda\in\Lambda} \cos(2\phi_\lambda) = 0 \quad and \quad \sum_{\lambda\in\Lambda} \sin(2\phi_\lambda) = 0 \tag{26}$$

hold.

The remainder of this section focuses on deriving and understanding certain physical consequences of Theorem 2. Of course these results apply also to Theorem 3.

Proposition 5. *The following actions do not affect the value of the Fisher information determinant: (i) changes in the true individual sensor-target ranges, i.e. moving a sensor from* $\mathbf{s}_i$ *to* $\mathbf{p} + k(\mathbf{s}_i - \mathbf{p})$ *for some* $k > 0$; *and (ii) reflecting a sensor about the emitter position, i.e. moving a sensor from* $\mathbf{s}_i$ *to* $2\mathbf{p} - \mathbf{s}_i$.

Part (ii) of the preceding proposition is illustrated in Figure 2. In particular, Figure 2 illustrates two scenarios obtained from each other by reflecting a particular sensor about the emitter position.

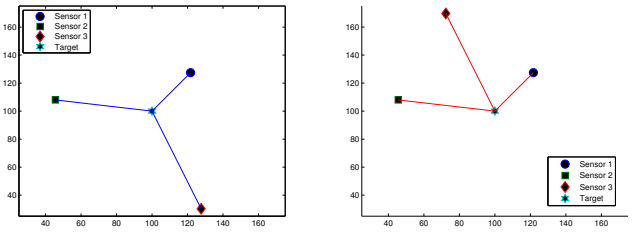


Fig. 2. This figure illustrates two scenarios obtained from each other by reflecting a particular sensor about the emitter position. This reflection does not affect the optimality of the sensor-target configuration.

Proposition 6. *Let the angle subtended at the target by two sensors* i *and* j *be denoted by* $\vartheta_{ij} = \vartheta_{ji}$. *One set of solutions to (25) is characterized by*

$$\vartheta_{ij} = \vartheta_{ji} = \frac{2}{n}\pi \tag{27}$$

for all adjacent sensor pairs $i, j \in \{1, \ldots, n \geq 4\}$ *with* $|j-i| = 1$ *or* $|j-i| = n-1$, *and then by a possible application of Proposition 5 on (27).*

Proof: The proof of this proposition is straightforward and involves verifying that (27) satisfies (25) of Theorem 2. Then, sensor reflections as detailed in Proposition 5 do not change the value of the determinant. Further details are omitted for brevity. ■

Corollary 2. *Consider the angles* $\vartheta_{ij} = \vartheta_{ji}$ *subtended at the target by two adjacent sensors where adjacency implies* $|j-i| = 1$ *or* $|j-i| = n-1$ *with* $i, j \in \{1, \ldots, n \geq 4\}$. *Then* $\vartheta_{ij} = \frac{2}{n}\pi$ *and* $\vartheta_{ij} = \frac{1}{n}\pi$ *are two separate solutions to (25).*

Consider now $n \geq 4$ sensors and denote the set of sensors by $\mathcal{V}$. Now assume that $\mathcal{V}$ can be partitioned into some arbitrary number m of subsets $\mathcal{B}_i$ such that $\mathcal{B}_i \cap \mathcal{B}_j = \emptyset$ and $|\mathcal{B}_i| \geq 2$, $\forall i, j \in \{1, \ldots, m\}$ with $i \neq j$. Then the following corollary follows directly from Theorem 2.

Corollary 3. *Given an arbitrary number* $n \geq 4$ *of sensors, then a solution to (25) can be obtained by arranging all subsets of sensors* $\mathcal{B}_i$, $\forall i \in \{1, \ldots, m\}$ *such that the condition of Theorem 2 is satisfied independently for those sensors in* $\mathcal{B}_i$, $\forall i \in \{1, \ldots, m\}$.

When $2 \leq n < 4$ the optimal sensor-target angular configuration for Doppler-based velocity estimation is unique up to sensor reflections about the target. When $n \geq 4$ there exists an infinite number of optimal configurations obtained by optimally placing disjoint subsets of sensors with cardinality ≥ 2. Corollary 3 is useful in forming optimal sensor configurations with an arbitrary number $n \geq 4$ of sensors. We demonstrate the flexibility this affords a designer in Figure 3.

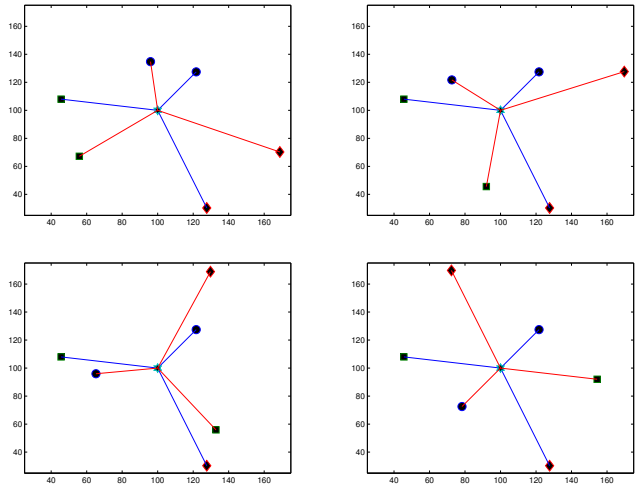


Fig. 3. Consider a mono-static or passive scenario involving $n = 6$ sensors. We can partition the sensors into two disjoint sets $\mathcal{B}_1 = \{1, 2, 3\}$ and $\mathcal{B}_2 = \{4, 5, 6\}$. This figure illustrates a number of scenarios where each disjoint subset of sensors is independently placed in an optimal configuration. Each scenario is optimal.

Figure 3 illustrates how disjoint subsets of sensors can be optimally and independently placed with respect to each other while still maintaining a globally optimal geometry in the sense that the angular positions obey (25) in Theorem 2.

V. Conclusion

In this paper we have outlined the problem of multi-static Doppler-based target position and velocity estimation. The Fisher information matrix was derived given a separate target illuminator and then given a target-based isotropic signal emission. Some remarks concerning the Cramer-Rao inequality and its relationship to the estimation problem were given. For the case of Doppler-based target velocity estimation we completely characterized the sensor-target geometry and provided a number of conditions on the optimal placement of the sensors and the transmitters.

Acknowledgment

The first author was supported by NICTA. NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

References

[1] L.R. Malling. Radio doppler effect for aircraft speed measurements. *Proceedings of the IRE*, 35(11):1357–1360, November 1947.

[2] E.J. Barlow. Doppler radar. *Proceedings of the IRE*, 37(4):340–355, April 1949.

[3] Walter R. Fried. Principles and performance analysis of doppler navigation systems. *IRE Transactions on Aeronautical and Navigational Electronics*, 4(4):176–196, December 1957.

[4] S.N. Salinger and J.J. Brandstatter. Application of recursive estimation and kalman filtering to doppler tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 19(4):585–592, July 1970.

[5] H.L. Van Trees. *Detection, Estimation and Modulation Theory*. John Wiley and Sons, Inc., New York, NY, 1968.

[6] A. Host-Madsen. On the existence of efficient estimators. *IEEE Transactions on Signal Processing*, 48(11):3028–3031, November 2000.

[7] S.C. Nardone, A.G. Lindgren, and K.F. Gong. Fundamental properties and performance of conventional bearings-only target motion analysis. *IEEE Transactions on Automatic Control*, AC-29(9):775–787, 1984.

[8] M. Gavish and A.J. Weiss. Performance analysis of bearing-only target location algorithms. *IEEE Transactions on Aerospace and Electronic Systems*, 28(3):817–827, 1992.

[9] Y.C. Eldar. Uniformly improving the Cramer-Rao bound and maximum-likelihood estimation. *IEEE Transactions on Signal Processing*, 54(8):2943–2956, August 2006.

[10] A.O. Hero, J.A. Fessler, and U. Usman. Exploring estimator bias-variance tradeoffs using the uniform CR bound. *IEEE Transactions on Signal Processing*, 44(8):2026–2041, August 1996.

[11] Y.C. Eldar. Minimum variance in biased estimation: Bounds and asymptotically optimal estimators. *IEEE Transactions on Signal Processing*, 52(7):1915–1930, July 2004.

[12] A.N. Bishop, B. Fidan, B.D.O. Anderson, K. Dogancay, and P.N. Pathirana. Optimality analysis of sensor-target localization geometries. *Automatica*, 46(3):479–492, March 2010.

2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Contractions for Consensus Processes*

J. Liu, A. S. Morse, B. D. O. Anderson, and C. Yu

***Abstract*— Many distributed control algorithms of current interest can be modeled by linear recursion equations of the form $x(t+1) = M(t)x(t),\ t \geq 1$ where each $M(t)$ is a real-valued "stochastic" or "doubly stochastic" matrix. Convergence of such recursions often reduces to deciding when the sequence of matrix products $M(1), M(2)M(1), M(3)M(2)M(1), \ldots$ converges. Certain types of stochastic and doubly stochastic matrices have the property that *any* sequence of products of such matrices of the form $S_1, S_2S_1, S_3S_2S_1, \ldots$ converges exponentially fast. We explicitly characterize the largest classes of stochastic and doubly stochastic matrices with positive diagonal entries which have these properties. The main goal of this paper is to find a "semi-norm" with respect to which matrices from these "convergability classes" are contractions. For any doubly stochastic matrix S such a semi-norm is identified and is shown to coincide with the second largest *singular* value of S.**

I. Introduction

Many distributed control algorithms of current interest can be modeled by linear recursion equations of the form

$$x(t+1) = M(t)x(t), \quad t \geq 1 \tag{1}$$

where each $M(t)$ is a real-valued "stochastic" or "doubly stochastic" matrix. Among these are consensus and flocking algorithms [2]–[8], distributed averaging algorithms [9]–[11], and certain types of gossiping algorithms [12]–[14]. Recursion equations like this have their roots in the literature on nonhomogeneous Markov chains [15]. While much is known at this point about conditions on the $M(t)$ for solutions to converge to a limit point, considerably less is known about the rates at which such solutions converge. There are classical concepts such as the coefficient of ergodicity [15] which are helpful in deriving convergence rates, but these are limited to only certain types of processes. The convergence rate question has been studied recently in [1], [11], [16], [17]. In [9], [11] convergence rate results are derived for distributed averaging algorithms. In [12] the question is addressed for probabilistic gossiping algorithms. A modified gossiping algorithm intended to speed up convergence is proposed in [18] without proof of correctness, but with convincing experimental results. The algorithm has recently been analyzed in [19]. Recent results concerning convergence rates appear in [13], [20]–[22] for periodic gossiping and in [1], [11], [23] for deterministic aperiodic gossiping.

Certain types of stochastic and doubly stochastic matrices have the property that *any* sequence of products of such matrices of the form $S_1, S_2S_1, S_3S_2S_1, \ldots$ converges exponentially fast. In Section II we explicitly characterize the largest classes of stochastic and doubly stochastic matrices with positive diagonal entries which have these properties. We call these classes "convergable". The main goal of this paper is to find a "semi-norm" with respect to which matrices from these convergability classes are contractions. The role played by semi-norms in characterizing convergence rates is explained in Section III. Three different types of semi-norms are considered. Each is compared to the well known coefficient of ergodicity which plays a central role in the study of convergence rates for nonhomogeneous Markov chains [15]. Somewhat surprisingly, for doubly stochastic matrices it turns out that a particular Euclidean semi-norm on $\mathbb{R}^{n\times n}$ has the required property - namely that in this semi-norm, any doubly stochastic matrix S in the convergability class of all doubly stochastic matrices is a contraction. This particular semi-norm turns out to be the second largest singular value of S.

A. Stochastic and Doubly Stochastic Matrices

The type of matrices typically encountered in a consensus process [4] modeled by (1) have only nonnegative entries and row sums all equal one. Matrices with these properties are called *stochastic*. *Doubly stochastic* matrices are stochastic matrices with the additional property that their column sums are also all equal to one. Doubly stochastic matrices are typically encountered when (1) represents a distributed averaging [9] or gossiping [12] process. It is easy to see that a nonnegative matrix S is stochastic if and only if $S\mathbf{1} = \mathbf{1}$ where $\mathbf{1} \in \mathbb{R}^n$ is a column vector whose entries are all ones. Similarly a nonnegative matrix S is doubly stochastic if and only if $S\mathbf{1} = \mathbf{1}$ and $S'\mathbf{1} = \mathbf{1}$. Using these characterizations it is easy to prove that the class of stochastic matrices in $\mathbb{R}^{n\times n}$ is compact and closed under multiplication as is the class of doubly stochastic matrices

Some of the research in this paper was presented at a workshop entitled "A Celebration of the Field of Systems and Control" which was held in Stockholm, Sweden, September, 2009; that research was subsequently published in shortened form in [1]. J. Liu and A. S. Morse are with Yale University, USA ({ji.liu, as.morse}@yale.edu). B. D. O. Anderson and C. Yu are with the Australian National University and National ICT Australia Ltd., Australia ({brian.anderson, brad.yu}@anu.edu.au). C. Yu is also with Shandong Computer Science Center, Jinan, China. The research of the first two authors is supported by the US Army Research Office, the US Air Force Office of Scientific Research, and the National Science Foundation. B. D. O. Anderson is supported by Australian Research Council's Discovery Project DP-110100538 and National ICT Australia-NICTA. NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. C. Yu is supported by the Australian Research Council through a Queen Elizabeth II Fellowship and DP-110100538 and Overseas Expert Program of Shandong Province. B. D. O. Anderson and C. Yu are also supported by the US Air Force Research laboratory grant number FA2386-10-1-4102.

*Proofs of some assertions in this paper are omitted due to space limitations; they are available from the first author upon request and will be given in a full length version of this paper.

in $\mathbb{R}^{n\times n}$. It is also true that the class of nonnegative matrices in $\mathbb{R}^{n\times n}$ with positive diagonal entries is closed under multiplication. Stochastic and doubly stochastic matrices with positive diagonal entries are commonly encountered in the study of consensus processes; positive diagonal entries greatly simplify convergence analysis.

Mathematically, reaching a consensus means that the state vector $x(t)$ appearing in (1) converges to a limit vector of the form $\alpha\mathbf{1}$ were α is a number depending on the initial value of x. This will always be the case if the infinite sequence of matrix products $M(1), M(2)M(1), M(3)M(2)M(1), \ldots$ converges to a matrix of the form $\mathbf{1}c$ in which case $\alpha = cx(1)$. It should be clear from what has just been stated that if the $M(t)$ are all doubly stochastic, then $c = \frac{1}{n}\mathbf{1}'$ which means that in this case α is the average of the values of the entries in $x(1)$. Thus to study convergence of the consensus process modeled by (1), it suffices to study the convergence of infinite sequences of products of stochastic and doubly stochastic matrices. Such sequences are closely related to what are called "nonhomogeneous Markov chains" for which there is a substantial literature [15]. Notwithstanding this, the following question remains. What determines the convergence rate of such sequences? This is the question which will be considered in the sequel. We begin with a few basic ideas.

B. Graph of a Stochastic Matrix

Many properties of a stochastic matrix can be usefully described in terms of an associated directed graph determined by the matrix. The *graph* of nonnegative matrix $M \in \mathbb{R}^{n\times n}$, written $\gamma(M)$, is a directed graph on n vertices with an arc from vertex i to vertex j just in case $m_{ji} \neq 0$; if (i, j) is such an arc, we say that i is a *neighbor* of j and that j is an *observer* of i. Thus $\gamma(M)$ is that directed graph whose adjacency matrix is the transpose of the matrix obtained by replacing all nonzero entries in M with ones.

C. Connectivity

There are various notions of connectivity which are useful in the study of the convergence of products of stochastic matrices. Perhaps the most familiar of these is the idea of "strong connectivity". A directed graph is *strongly connected* if there is a directed path between each pair of distinct vertices. A directed graph is *weakly connected* if there is an undirected path between each pair of distinct vertices. There are other notions of connectivity which are also useful in this context. To define several of them, let us agree to call a vertex i of a directed graph $\mathbb{G}$, a *root* of $\mathbb{G}$ if for each other vertex j of $\mathbb{G}$, there is a directed path from i to j. Thus i is a root of $\mathbb{G}$ if it is the root of a directed spanning tree of $\mathbb{G}$. We will say that $\mathbb{G}$ is *rooted at* i if i is in fact a root. Thus $\mathbb{G}$ is rooted at i just in case each other vertex of $\mathbb{G}$ is *reachable* from vertex i along a directed path within the graph. $\mathbb{G}$ is *strongly rooted at* i if each other vertex of $\mathbb{G}$ is reachable from vertex i along a directed path of length 1. Thus $\mathbb{G}$ is strongly rooted at i if i is a neighbor of every other vertex in the graph. By a *rooted graph* is meant a directed graph which possesses at least one root. A *strongly rooted graph* is a graph which has at least one vertex at which it is strongly rooted. Note that a nonnegative matrix $M \in \mathbb{R}^{n\times n}$ has a strongly rooted graph if and only if it has a positive column. Note that every strongly connected graph is rooted and every rooted graph is weakly connected. The converse statements are false. In particular there are weakly connected graphs which are not rooted and rooted graphs which are not strongly connected.

D. Composition

Since we will be interested in products of stochastic matrices, we will be interested in graphs of such products and how they are related to the graphs of the matrices comprising the products. For this we need the idea of "composition" of graphs. Let $\mathbb{G}_p$ and $\mathbb{G}_q$ be two directed graphs with vertex set $\mathcal{V}$. By the *composition* of $\mathbb{G}_p$ with $\mathbb{G}_q$, written $\mathbb{G}_q \circ \mathbb{G}_p$, is meant the directed graph with vertex set $\mathcal{V}$ and arc set defined in such a way so that (i, j) is an arc of the composition just in case there is a vertex k such that (i, k) is an arc of $\mathbb{G}_p$ and (k, j) is an arc of $\mathbb{G}_q$. Thus (i, j) is an arc in $\mathbb{G}_q \circ \mathbb{G}_p$ if and only if i has an observer in $\mathbb{G}_p$ which is also a neighbor of j in $\mathbb{G}_q$. Note that composition is an associative binary operation; because of this, the definition extends unambiguously to any finite sequence of directed graphs $\mathbb{G}_1, \mathbb{G}_2, \ldots, \mathbb{G}_k$ with the same vertex set.

Composition and matrix multiplication are closely related. In particular, the graph of the product of two nonnegative matrices $M_1, M_2 \in \mathbb{R}^{n\times n}$ is equal to the composition of the graphs of the two matrices comprising the product. In other words, $\gamma(M_2M_1) = \gamma(M_2) \circ \gamma(M_1)$.

If we focus exclusively on graphs with self-arcs at all vertices, more can be said. In this case the definition of composition implies that the arcs of both $\mathbb{G}_p$ and $\mathbb{G}_q$ are arcs of $\mathbb{G}_q \circ \mathbb{G}_p$; the converse is false. The definition of composition also implies that if $\mathbb{G}_p$ has a directed path from i to k and $\mathbb{G}_q$ has a directed path from k to j, then $\mathbb{G}_q \circ \mathbb{G}_p$ has a directed path from i to j. These implications are consequences of the requirement that the vertices of the graphs in question have self-arcs at all vertices. It is worth emphasizing that the union of the arc sets of a sequence of graphs $\mathbb{G}_1, \mathbb{G}_2, \ldots, \mathbb{G}_k$ with self-arcs must be contained in the arc set of their composition. However the converse is not true in general and it is for this reason that composition rather than union proves to be the more useful concept for our purposes.

II. Convergability

It is of obvious interest to have a clear understanding of what kinds of stochastic matrices within an infinite product guarantee that the infinite product converges. There are many ways to address this issue and many existing results. Here we focus on just one issue.

Let $\mathcal{S}$ denote the set of all stochastic matrices in $\mathbb{R}^{n\times n}$ with positive diagonal entries. Call a compact subset $\mathcal{M} \subset \mathcal{S}$ *convergable* if for each infinite sequence of matrices $M_1, M_2, M_3, \ldots$ from $\mathcal{M}$, the sequence of products

$M_1, M_2M_1, M_3M_2M_1, \ldots$ converges exponentially fast to a matrix of the form $\mathbf{1}c$. Convergability can be characterized as follows.

Theorem 1: Let $\mathcal{R}$ denote the set of all matrices in $\mathcal{S}$ with rooted graphs. Then a compact subset $\mathcal{M} \subset \mathcal{S}$ is convergable if and only if $\mathcal{M} \subset \mathcal{R}$.

The theorem implies that $\mathcal{R}$ is the largest subset of $n \times n$ stochastic matrices with positive diagonal entries whose compact subsets are all convergable. $\mathcal{R}$ itself is not convergable because it is not closed and thus not compact.

Proof of Theorem 1: The fact that any compact subset of $\mathcal{R}$ is convergable is more or less well known from the work reported in [24]; the statement also follows from Proposition 11 of [25]. To prove the converse, suppose that $\mathcal{M} \subset \mathcal{S}$ is convergable. Then by continuity, every sufficiently long product of matrices from $\mathcal{M}$ must be a matrix with a positive column. Therefore, the graph of every sufficiently long product of matrices from $\mathcal{M}$ must be strongly rooted. It follows from Proposition 5 of [25] that $\mathcal{M}$ must be a subset of $\mathcal{R}$. ■

Although doubly stochastic matrices are stochastic, convergability for classes of doubly stochastic matrices has a different characterization than it does for classes of stochastic matrices. Let $\mathcal{D}$ denote the set of all doubly stochastic matrices in $\mathcal{S}$. In the sequel we will prove the following theorem.

Theorem 2: Let $\mathcal{W}$ denote the set of all matrices in $\mathcal{D}$ with strongly connected graphs. Then a compact subset $\mathcal{M} \subset \mathcal{D}$ is convergable if and only if $\mathcal{M} \subset \mathcal{W}$.

The theorem implies that $\mathcal{W}$ is the largest subset of $n \times n$ doubly stochastic matrices with positive diagonal entries whose compact subsets are all convergable. Like $\mathcal{R}$, $\mathcal{W}$ is not convergable because it is not compact. Results which more or less imply the sufficiency of strong connectivity can be found in [24] and elsewhere. Note that sufficiency is also implied by Theorem 1 because doubly stochastic matrices with strongly connected graphs are stochastic matrices with rooted graphs. It remains therefore, to prove the necessity of Theorem 2. This will be done in the sequel.

An interesting set of stochastic matrices in $\mathcal{S}$ whose compact subsets are known to be convergable, is the set of all "scrambling matrices". A matrix $S \in \mathcal{S}$ is *scrambling* if for each distinct pair of integers i and j, there is a column k of S for which s_{ik} and s_{jk} are both nonzero [15]. In graph theoretic terms S is a scrambling matrix just in case its graph is "neighbor shared" where by *neighbor shared* we mean that each distinct pair of vertices in the graph share a common neighbor [25]. Convergability of compact subsets of scrambling matrices is tied up with the concept of the coefficient of ergodicity [15] which for a given stochastic matrix $S \in \mathcal{S}$ is defined by the formula

$$\tau(S) = \frac{1}{2} \max_{i,j} \sum_{k=1}^{n} |s_{ik} - s_{jk}| \tag{2}$$

It is known that $0 \leq \tau(S) \leq 1$ for all $S \in \mathcal{S}$ and that

$$\tau(S) < 1 \tag{3}$$

if and only if S is a scrambling matrix. It is also known that

$$\tau(S_2S_1) \leq \tau(S_2)\tau(S_1), \quad S_1, S_2 \in \mathcal{S} \tag{4}$$

It can be shown that (3) and (4) are sufficient conditions to ensure that any compact subset of scrambling matrices is convergable. But $\tau(\cdot)$ has another role. It provides a worst case convergence rate for any infinite product of scrambling matrices from a given compact set $\mathcal{C} \subset \mathcal{S}$. In particular, it can be easily shown that as $i \to \infty$, any product $S_iS_{i-1} \cdots S_2S_1$ of scrambling matrices $S_i \in \mathcal{C}$ converges to a matrix of the form $\mathbf{1}c$ as fast as λ^i converges to zero where

$$\lambda = \max_{S \in \mathcal{C}} \tau(S)$$

This preceding discussion suggests the following question. Can analogs of the coefficient of ergodicity satisfying formulas like (3) and (4) be found for the set of stochastic matrices with rooted graphs or perhaps for the set of doubly stochastic matrices with strongly connected graphs? In the sequel we will provide a partial answer to this question for the case of stochastic matrices and a complete answer for the case of doubly stochastic matrices. Our approach will be to appeal to certain types of semi-norms of stochastic matrices.

III. Semi-norms

Let $||\cdot||_p$ be the induced p-norm on $\mathbb{R}^{m\times n}$. In this paper we will be primarily interested in the cases $p = 1, 2, \infty$. Note that for a nonnegative matrix A

$$\begin{aligned} ||A||_1 &= \max \text{column sum } A \\ ||A||_2 &= \sqrt{\mu(A'A)} \\ ||A||_\infty &= \max \text{row sum } A \end{aligned}$$

where $\mu(A'A)$ is the largest eigenvalue of $A'A$; that is, the square of the largest singular value of A. For any integer $p > 0$ and matrix $M \in \mathbb{R}^{m\times n}$ define

$$|M|_p = \min_{c \in \mathbb{R}^{1\times n}} ||M - \mathbf{1}c||_p$$

As defined, $|\cdot|_p$ is nonnegative and $|M|_p \leq ||M||_p$; clearly $|\mu M|_p = |\mu||M|_p$ for all real numbers μ so $|\cdot|_p$ is "positively homogeneous" [26]. Let M_1 and M_2 be matrices in $\mathbb{R}^{m\times n}$ and let c_0, c_1, and c_2 denote values of c which minimize $||M_1+M_2-\mathbf{1}c||_p$, $||M_1-\mathbf{1}c||_p$, and $||M_2-\mathbf{1}c||_p$ respectively. Note that

$$\begin{aligned} |M_1 + M_2|_p &= ||M_1 + M_2 - \mathbf{1}c_0||_p \\ &\leq ||M_1 + M_2 - \mathbf{1}(c_1 + c_2)||_p \\ &\leq ||M_1 - \mathbf{1}c_1||_p + ||M_2 - \mathbf{1}c_2||_p \\ &= |M_1|_p + |M_2|_p \end{aligned}$$

Thus the triangle inequality holds. These properties mean that $|\cdot|_p$ is a *semi-norm*. $|\cdot|_p$ behaves much like a norm. For example, if N is a submatrix of M, then $|N|_p \leq |M|_p$. However $|\cdot|_p$ is not a norm because $|M|_p = 0$ does not imply $M = 0$; rather it implies that $M = \mathbf{1}c$ for some row vector c which minimizes $||M-\mathbf{1}c||_p$. For our purposes, $|\cdot|_p$ has a particularly important property:

Lemma 1: Suppose $\mathcal{M}$ is a subset of $\mathbb{R}^{n\times n}$ such that $M\mathbf{1} = \mathbf{1}$ for all $M \in \mathcal{M}$. Then

$$|M_2M_1|_p \leq |M_2|_p|M_1|_p \tag{5}$$

Proof of Lemma 1: Let c_0, c_1, and c_2 denote values of c which minimize $||M_2M_1 - \mathbf{1}c||_p$, $||M_1 - \mathbf{1}c||_p$, and $||M_2 - \mathbf{1}c||_p$ respectively. Then

$$\begin{aligned}
|M_2M_1|_p &= ||M_2M_1 - \mathbf{1}c_0||_p \\
&\leq ||M_2M_1 - \mathbf{1}(c_2M_1 + c_1 - c_2\mathbf{1}c_1)||_p \\
&= ||M_2M_1 - \mathbf{1}c_2M_1 - M_2\mathbf{1}c_1 + \mathbf{1}c_2\mathbf{1}c_1)||_p \\
&= ||(M_2 - \mathbf{1}c_2)(M_1 - \mathbf{1}c_1)||_p \\
&\leq ||(M_2 - \mathbf{1}c_2)||_p||(M_1 - \mathbf{1}c_1)||_p \\
&= |M_2|_p|M_1|_p
\end{aligned}$$

Thus (5) is true. ■

We say that $M \in \mathbb{R}^{n\times n}$ is *semi-contractive* in the p-norm if $|M|_p < 1$. In view of Lemma 1, the product of semi-contractive matrices in $\mathcal{M}$ is thus semi-contractive. The importance of these ideas lies in the following fact.

Proposition 1: Suppose $\mathcal{M}$ is a subset of $\mathbb{R}^{n\times n}$ such that $M\mathbf{1} = \mathbf{1}$ for all $M \in \mathcal{M}$. Let p be fixed and let $\bar{\mathcal{M}}$ be a compact set of semi-contractive matrices in $\mathcal{M}$. Let

$$\lambda = \sup_{\bar{\mathcal{M}}} |M|_p$$

Then for each infinite sequence of matrices $M_i \in \bar{\mathcal{M}}$, $i \in \{1, 2, \ldots\}$, the matrix product $M_iM_{i-1}\cdots M_1$ converges as $i \to \infty$ as fast as λ^i converges to zero, to a rank one matrix of the form $\mathbf{1}\bar{c}$.

A. The case $p = 1$

We now consider in more detail the case when $p = 1$. For this case it is possible to derive an explicit formula for the semi-norm $|M|_1$ of a nonnegative matrix $M \in \mathbb{R}^{n\times n}$.

Proposition 2: Let q be the unique integer quotient of n divided by 2. Let $M \in \mathbb{R}^{n\times n}$ be a nonnegative matrix. Then

$$|M|_1 = \max_{j\in\{1,2,\ldots,n\}} \left\{ \sum_{i\in\mathcal{L}_j} m_{ij} - \sum_{i\in\mathcal{S}_j} m_{ij} \right\}$$

where $\mathcal{L}_j$ and $\mathcal{S}_j$ are respectively the row indices of the q largest and q smallest entries in the jth column of M.

Consider now the case when M is a doubly stochastic matrix S, more can be said:

Theorem 3: Let q be the unique integer quotient of n divided by 2. Let $S \in \mathbb{R}^{n\times n}$ be a doubly stochastic matrix. Then $|S| \leq 1$. Moreover S is a semi-contraction in the one-norm if and only if the number of nonzero entries in each column of S exceeds q.

Note that the doubly stochastic matrix

$$S = \begin{bmatrix} .5 & .125 & .125 & .125 & .125 \\ .5 & .125 & .125 & .125 & .125 \\ 0 & .25 & .25 & .25 & .25 \\ 0 & .25 & .25 & .25 & .25 \\ 0 & .25 & .25 & .25 & .25 \end{bmatrix}$$

has a strongly connected graph but is not a semi-contractions for $p = 1$. Thus this particular semi-norm is not as useful as we would like because there are matrices in $\mathcal{W}$ which are not semi-contractions for $p = 1$.

It is possible to compare this semi-norm with the coefficient of ergodicity. Observe that while the preceding matrix is not a semi-contraction it is a scrambling matrix. Thus for this example, $\tau(S) < |S|_1 = 1$. On the other hand there are also doubly stochastic matrices which are semi-contractions but which are not scrambling matrices. An example of this is the matrix

$$S = \begin{bmatrix} .5 & 0 & 0 & 0 & .5 & 0 \\ 0 & .5 & 0 & 0 & 0 & .5 \\ .125 & .125 & .25 & .25 & .125 & .125 \\ .125 & .125 & .25 & .25 & .125 & .125 \\ .125 & .125 & .25 & .25 & .125 & .125 \\ .125 & .125 & .25 & .25 & .125 & .125 \end{bmatrix}$$

Thus for this example, $|S|_1 < \tau(S) = 1$, which means that there are situations when it may be more advantageous to use the semi-norm $|\cdot|_1$ to compute convergence rates than to appeal to the coefficient of ergodicity.

B. The case $p = \infty$

Note that in this case $|S|_\infty \leq 1$ for any stochastic matrix because $|S|_\infty \leq ||S||_\infty = 1$. Although not at all obvious, it turns out that $|S|_\infty$ equals the coefficient of ergodicity discussed earlier and defined by (2). This is an immediate consequence of Proposition 3 which is stated below. Unfortunately, the last example in the preceding subsection shows that there are doubly stochastic matrices with strongly connected graphs which are not scrambling matrices. Thus this particular semi-norm is also not as useful as we might hope for.

Proposition 3: Let $A \in \mathbb{R}^{n\times n}$ be a nonnegative matrix. Then

$$|A|_\infty = \frac{1}{2}\max_{i,j} \sum_{k=1}^{n} |a_{ik} - a_{jk}|$$

C. The case $p = 2$

For the case when $p = 2$ it is also possible to derive an explicit formula for the semi-norm $|M|_2$ of a nonnegative matrix $M \in \mathbb{R}^{n\times n}$. Towards this end note that for any $x \in \mathbb{R}^n$, the function $g(x, c) = x'(M - \mathbf{1}c)'(M - \mathbf{1}c)x$ attains its minimum with respect to c at $\frac{1}{n}\mathbf{1}'M$. This implies that

$$|M|_2 = ||PM||_2 = \sqrt{\mu\{M'P'PM\}}$$

where $P = I - \frac{1}{n}\mathbf{1}\mathbf{1}'$ and, for any symmetric matrix T, $\mu\{T\}$ is the largest eigenvalue of T. We are led to the following result.

Proposition 4: Let $M \in \mathbb{R}^{n\times n}$ be a nonnegative matrix. Then $|M|_2$ is the largest singular value of the matrix PM where P is the orthogonal projection on the orthogonal complement of the span of $\mathbf{1}$.

Now suppose that M is a doubly stochastic matrix S. Then $S'S$ is also doubly stochastic and $\mathbf{1}'S = \mathbf{1}'$. The latter and Proposition 4 imply that

$$|S|_2 = \sqrt{\mu\left\{S'S - \frac{1}{n}\mathbf{1}\mathbf{1}'\right\}} \tag{6}$$

More can be said:

Lemma 2: If S is doubly stochastic, then $\mu\{S'S - \frac{1}{n}\mathbf{1}\mathbf{1}'\}$ is the second largest eigenvalue of $S'S$.

Proof of Lemma 2: Since $S'S$ is symmetric it has orthogonal eigenvectors one of which is $\mathbf{1}$. Let $\mathbf{1}, x_2, \ldots, x_n$ be such a set of eigenvectors with eigenvalues $1, \lambda_2, \ldots, \lambda_n$. Then $S'S\mathbf{1} = \mathbf{1}$ and $S'Sx_i = \lambda_i x_i,\ i \in \{2, 3, \ldots, n\}$. Clearly $(S'S - \frac{1}{n}\mathbf{1}\mathbf{1}')\mathbf{1} = 0$ and $(S'S - \frac{1}{n}\mathbf{1}\mathbf{1}')x_i = \lambda_i x_i,\ i \in \{2, 3, \ldots, n\}$. Since 1 is the largest eigenvalue of $S'S$ it must therefore be true that the second largest eigenvalue $S'S$ is the largest eigenvalue of $S'S - \frac{1}{n}\mathbf{1}\mathbf{1}'$. ■

We summarize:

Theorem 4: For $p = 2$ the semi-norm of a doubly stochastic matrix S is the second largest singular value of S.

There is another way to think about what this theorem implies. Prompted by the work in [9] and [11], suppose one wants to measure in the sense of a 2-norm $||\cdot||$, how much closer an n-vector x gets to the average vector $z = \frac{1}{n}\mathbf{1}\mathbf{1}'x$ when it is multiplied by a doubly stochastic matrix S. In other words how does the norm $||Sx - z||$ compare with $||x - z||$? To address this question, note first that $x - z \in \mathcal{O}$ where $\mathcal{O}$ is the orthogonal complement of the span of $\mathbf{1}$. Note next that

$$||Sx - z||^2 = ||S(x - z)||^2 \leq \left(\sup_{y \in \mathcal{O}} \frac{y'S'Sy}{y'y}\right) ||x - z||^2$$

But $\sup_{y \in \mathcal{O}} \frac{y'S'Sy}{y'y}$ is the second largest eigenvalue of $S'S$ which in turn is the square of the second largest singular value of S. In other words, $||Sx - z|| \leq |S|_2 ||x - z||$. Thus Sx is always as close to the average vector z as x is and is even closer if $|S|_2$ is a contraction.

In the light of Theorem 4, we are now in a position to characterize in graph theoretic terms those doubly stochastic matrices with positive diagonal entries which are semi-contractions for $p = 2$.

Theorem 5: Let S be a doubly stochastic matrix with positive diagonal entries. Then $|S|_2 \leq 1$. Moreover S is a semi-contraction in the 2-norm if and only if the graph of S is strongly connected.

To prove this theorem we need several concepts and results. Let $\mathbb{G}$ denote a directed graph and write $\mathbb{G}'$ for that graph which results when the arcs in $\mathbb{G}$ are reversed; i.e., the *dual graph*. Call a graph *symmetric* if it is equal to its dual. Note that in the case of a symmetric graph, the three properties of being rooted, strongly connected, and weakly connected are equivalent. Note also that if $\mathbb{G}$ is the graph of a nonnegative matrix M with positive diagonal entries, then $\mathbb{G}'$ is the graph of M' and $\mathbb{G}' \circ \mathbb{G}$ is the graph of $M'M$.

Lemma 3: A directed graph $\mathbb{G}$ with self-arcs at all vertices is weakly connected if and only if $\mathbb{G}' \circ \mathbb{G}$ is strongly connected.

Lemma 4: Let T be a stochastic matrix with positive diagonal entries. If T has a strongly connected graph, then the magnitude of its second largest eigenvalue is less than one. If, on the other hand, the magnitude of the second largest eigenvalue of T is less than one, then the graph of T is weakly connected.

Lemma 5: The graph $\mathbb{G}$ of a doubly stochastic matrix D is strongly connected if and only if it is weakly connected.[1]

The proof of Lemma 5 which follows is based on ideas from [15] and [27]. Let $\mathbb{G}$ be a directed graph with vertex set $\mathcal{V} = \{1, 2, \ldots, n\}$. Call a vertex j is *reachable* from i if either $j = i$ or if there is a directed path from i to j. Call a vertex i *essential* if i is reachable from all vertices which are reachable from i.

Lemma 6: Every directed graph has at least one essential vertex.

To proceed, let us say that vertices i and j are *mutually reachable* if each is reachable from the other. Mutual reachability is clearly an equivalence relation on $\mathcal{V}$ which partitions $\mathcal{V}$ into the disjoint union of a finite number of equivalence classes. Note that if i is an essential vertex of $\mathbb{G}$, then every vertex in the equivalence class of i is also essential. Thus every directed graph possesses at least one mutually reachable equivalence class whose members are all essential.

Proof of Lemma 5: Strong connectivity clearly implies weak connectivity. We prove the converse. Suppose $\mathbb{G}$ is weakly connected. In view of the proceeding, $\mathbb{G}$ has at least one mutually reachable equivalence class $\mathcal{E}$ whose members are all essential. If $\mathcal{E} = \mathcal{V}$, then $\mathbb{G}$ is obviously strongly connected. Thus to prove the lemma, it is enough to show that $\mathcal{E} = \mathcal{V}$. Suppose the contrary, namely that $\mathcal{E} = \{i_1, i_2, \ldots, i_m\}$ is a strictly proper subset of $\mathcal{V}$. Let π be any permutation map for which $\pi(i_j) = j,\ j \in \{1, 2, \ldots, m\}$ and let P be the corresponding permutation matrix. Then clearly

$$P'DP = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

and $P'DP$ is doubly stochastic. Since $P'DP$ is doubly stochastic, the column sums of A must all equal one as must the row sums of the submatrix $[\,A \quad B\,]$. But the transformation $D \longmapsto P'DP$ corresponds to a relabeling of the vertices of $\mathbb{G}$, so the graph of $P'DP$ must also be weakly connected. Thus means that B cannot be the zero matrix. Therefore the sum of the row sums of A must be less than m. But this contradicts the fact that the sum of the column sums of A equals m. Therefore $\mathcal{E} = \mathcal{V}$. ■

Proof of Theorem 5: Let S be a doubly stochastic matrix with positive diagonal entries. Then 1 is the largest singular value of S because $S'S$ is doubly stochastic. From this and Theorem 4 it follows that $|S|_2 \leq 1$.

Suppose S is a semi-contraction. Then in view of Theorem 4, the second largest eigenvalue of $S'S$ is less than 1. Thus by Lemma 4, the graph of $S'S$ is weakly connected. But

[1]It is clear that strong connectivity of $\mathbb{G}$ implies weak connectivity of $\mathbb{G}$. The converse was conjectured by John Tsitsiklis in a private communication.

$S'S$ is symmetric so its graph must be strongly connected. Therefore by Lemma 3, the graph of S is weakly connected. In view of Lemma 5, the graph of S is strongly connected.

Now suppose that the graph of S is strongly connected. Then S is weakly connected so the graph of $S'S$ is strongly connected because of Lemma 3. Thus by Lemma 4, the magnitude of the second largest eigenvalue of $S'S$ is less than 1. From this and Theorem 4 it follows that S is a semi-contraction. ■

Proof of Theorem 2: Let $\mathcal{M}$ be any compact subset of $\mathcal{W}$. In view of Theorem 5, each matrix in $\mathcal{M}$ is a semi-contraction in the two-norm. From this and Proposition 1, it follows that $\mathcal{M}$ is convergable.

Now suppose that $\mathcal{M}$ is convergable and let S be a matrix in $\mathcal{M}$. Then S^i converges to a matrix of the form $\mathbf{1}c$ as $i \to \infty$. This means that the second largest eigenvalue of S must be less than 1 in magnitude. Thus by Lemma 4, S must have a weakly connected graph. By Lemma 5, the graph of S must be strongly connected. ■

The importance of Theorem 5 lies in the fact that the matrices in every convergable set of doubly stochastic matrices are contractions in the 2-norm. In view of Proposition 1, this enables one to immediately compute a rate of convergence for any infinite product of matrices from any given convergable set. The coefficient of ergodicity mentioned earlier does not have this property. If it did, then every doubly stochastic matrix with a strongly connected graph would have to be a scrambling matrix. The following counterexample shows that this is not the case:

$$S = \begin{bmatrix} .5 & .25 & 0 & 0 & 0 & .25 \\ .25 & .5 & 0 & 0 & 0 & .25 \\ 0 & 0 & .5 & .5 & 0 & 0 \\ 0 & 0 & .5 & .25 & 0 & .25 \\ 0 & 0 & 0 & 0 & .875 & .125 \\ .25 & .25 & 0 & .25 & .125 & .125 \end{bmatrix}$$

In particular, S is a doubly stochastic matrix with a strongly connected graph but it is not a scrambling matrix.

IV. Concluding Remarks

In this paper we have identified the largest "alphabets" of stochastic and doubly stochastic matrices with positive diagonal entries whose "words" converge exponentially fast as word length increases. In the case of double stochastic matrices, each matrix in the corresponding alphabet is shown to be a semi-contraction in the two-norm. In the case of stochastic matrices which are not doubly stochastic, we were not similarly successful and the problem of discovering a suitable semi-contraction for this case remains unresolved.

References

[1] J. Liu, A. S. Morse, B. D. O. Anderson, and C. Yu. The contraction coefficient of a complete gossip sequence. In *Three Decades of Progress in Control Sciences*, pages 275–290. Springer, 2010.

[2] J. N. Tsitsiklis. *Problems in decentralized decision making and computation*. PhD thesis, Department of Electrical Engineering and Computer Science, MIT, 1984.

[3] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, 31(9):803–812, 1986.

[4] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003.

[5] R. Olfati-Saber and R. M. Murray. Consensus seeking in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.

[6] V. D. Blondel, J. M. Hendrichkx, A. Olshevsky, and J. N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 2996–3000, 2005.

[7] L. Moreau. Stability of multi-agent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50(2):169–182, 2005.

[8] W. Ren and R. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, 50(5):655–661, 2005.

[9] L. Xiao and S. Boyd. Fast linear iterations for distributed averaging. *Systems and Control Letters*, 53(1):65–78, 2004.

[10] L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *Proceedings of the 4th International Conference on Information Processing in Sensor Networks*, pages 63–70, 2005.

[11] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis. On distributed averaging algorithms and quantization effects. *IEEE Transactions on Automatic Control*, 54(11):2506–2517, 2009.

[12] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

[13] B. D. O. Anderson, C. Yu, and A. S. Morse. Convergence of periodic gossiping algorithms. In *Perspectives in Mathematical System Theory, Control, and Signal Processing*, pages 127–138. Springer, 2010.

[14] J. Liu, S. Mou, A. S. Morse, B. D. O. Anderson, and C. Yu. Request-based gossiping. In *Proceedings of the 50th IEEE Conference on Decision and Control*, 2011. to be published.

[15] E. Seneta. *Non-negative matrices and Markov chains*. Sringer, 2006.

[16] M. Cao, D. A. Spielman, and A. S. Morse. A lower bound on convergence of a distributed network consencus algorithm. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 2356–2361, 2005.

[17] A. Olshevsky and J. N. Tsitsiklis. Convergence rates in distributed consensus and averaging. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 3387–3392, 2006.

[18] M. Cao, D. A. Spielman, and E. M. Yeh. Accelerated gossip algorithms for distributed computation. In *Proceedings of the 44th Annual Allerton Conference on Communication, Control, and Computing*, pages 952–959, 2006.

[19] J. Liu, B. D. O. Anderson, M. Cao, and A. S. Morse. Analysis of accelerated gossip algorithms. In *Proceedings of the 48th IEEE Conference on Decision and Control*, pages 871–876, 2009.

[20] S. Mou, C. Yu, B. D. O. Anderson, and A. S. Morse. Deterministic gossiping with a periodic protocol. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 5787–5791, 2010.

[21] F. He, A. S. Morse, J. Liu, and S. Mou. Periodic gossiping. In *Proceedings of the 18th IFAC World Congress*, 2011. to be published.

[22] C. Yu, B. D. O. Anderson, S. Mou, J. Liu, F. He, and A. S. Morse. Gossiping periodically. *IEEE Transactions on Automatic Control*, 2011. submitted.

[23] A. Olshevsky and J. N. Tsitsiklis. Convergence speed in distributed consensus and averaging. *SIAM Journal Control and Optimization*, 48(1):33–55, 2009.

[24] J. Wolfowitz. Products of indecomposable, aperiodic, stochastic matrices. *Proceedings of the American Mathematical Society*, 15(5):733–737, 1963.

[25] M. Cao, A. S. Morse, and B. D. O. Anderson. Reaching a consensus in a dynamically changing environment: a graphical approach. *SIAM Journal on Control and Optimization*, 47(2):575–600, 2008.

[26] R. C. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[27] R. G. Gallager. *Discrete Stochastic Processes*. Kluwer Academic Publishers, 1996.

2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Request-Based Gossiping

J. Liu, S. Mou, A. S. Morse, B. D. O. Anderson, and C. Yu

Abstract— **By the distributed averaging problem is meant the problem of computing the average value of a set of numbers possessed by the agents in a distributed network using only communication between neighboring agents. Gossiping is a well-known approach to the problem which seeks to iteratively arrive at a solution by allowing each agent to interchange information with at most one neighbor at each iterative step. Crafting a gossiping protocol which accomplishes this is challenging because gossiping is an inherently collaborative process which can lead to deadlock unless careful precautions are taken to ensure that it does not. In this paper we present three gossiping protocols. We show by example that the first can deadlock. While the second cannot, it requires a degree of network-wide coordination which may not be possible to secure in some applications. The third protocol uses only local information, is guaranteed to avoid deadlock, and requires fewer transmissions per iteration than standard broadcast-based distributed averaging protocols.**

I. Introduction

There has been considerable interest recently in developing algorithms for distributing information among the members of a group of sensors or mobile autonomous agents via local interactions. Notable among these are those algorithms intended to cause such a group to reach a consensus in a distributed manner [1]–[6]. One particular type of consensus processes which has received much attention lately is called distributed averaging [7]. In its simplest form, distributed averaging deals with a network of $n > 1$ *agents* and the constraint that each agent i is able to communicate only with certain other agents called agent i's *neighbors*. Neighbor relations are described by a simple, connected graph $\mathbb{A}$ in which vertices correspond to agents and edges indicate neighbor relations. Initially, each agent has or acquires a real number y_i which might be a measured temperature or something similar. The *distributed averaging problem* is to devise a protocol which will enable each agent to compute the average $y_{\rm avg} = \frac{1}{n}\sum_{i=1}^{n} y_i$ using only information acquired from its neighbors. There are many variants of this problem. For example, instead of real numbers, the y_i may be integer-valued [8]. Another variant assumes that the edges of $\mathbb{A}$ change over time [9]. This paper considers the case when the y_i are real and $\mathbb{A}$ does not depend on time.

As noted in [7], the distributed averaging problem can be solved, in principle, by "flooding"; that is, by propagating across the network over time the values of all of the y_i. Armed with knowledge of all of these values, each agent is thus able to compute $y_{\rm avg}$. A more sophisticated approach to the problem is for each agent to use a linear iterative update rule of the general form

$$x_i(t+1) = w_{ii}x_i(t) + \sum_{j\in\mathcal{N}_i} w_{ij}x_j(t), \qquad x_i(0) = y_i$$

where t is a discrete time index, $x_i(t)$ is agent i's current estimate of $y_{\rm avg}$, the w_{ij} are real-valued weights, and $\mathcal{N}_i$ is the set of labels of the neighbors of agent i. In [7] several methods are proposed for choosing the w_{ij}. One particular choice, which defines what has come to be known as the Metropolis algorithm, requires only local information to define the w_{ij}. Algorithms of this type, which require each agent to communicate with all of its neighbors on each iteration, are sometimes called *broadcast* algorithms.

An alternative approach to distributed averaging, which typically does not involve broadcasting, exploits a form of "gossiping" [10] specifically tailored to the distributed averaging problem. The idea of gossiping is very simple. A pair of neighbors with labels i and j are said to *gossip* at time t if both $x_i(t+1)$ and $x_j(t+1)$ are set equal to the average of $x_i(t)$ and $x_j(t)$. Each agent is allowed to gossip with at most one neighbor at one time. Under appropriate assumptions, algorithms which possess this simple property can be shown to solve the distributed averaging problem. Gossiping algorithms do not necessarily involve broadcasting and thus have the potential to require less transmissions per iteration than broadcast algorithms. Of course one would not expect gossip algorithms to converge as fast as broadcast algorithms.

The actual sequence of gossip pairs which occurs during a specific gossip process might be determined either probabilistically [10] or deterministically [11], [12], depending on the problem of interest. Deterministic gossiping protocols are intended to guarantee that under all conditions, a consensus will be achieved asymptotically whereas probabilistic protocols aim at achieving consensus asymptotically with probability one. Both approaches have merit. The probabilistic approach is typically somewhat easier both in terms

The authors thank Alex Olshevsky for useful discussions which have contributed to this work. The research of the first three authors is supported by the US Army Research Office, the US Air Force Office of Scientific Research, and the National Science Foundation. B. D. O. Anderson is supported by Australian Research Council's Discovery Project DP-110100538 and National ICT Australia-NICTA. NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. C. Yu is supported by the Australian Research Council through a Queen Elizabeth II Fellowship and DP-110100538 and Overseas Expert Program of Shandong Province. B. D. O. Anderson and C. Yu are also supported by the US Air Force Research laboratory grant number FA2386-10-1-4102.

J. Liu, S. Mou, and A. S. Morse are with Yale University, USA ({ji.liu, shaoshuai.mou, as.morse}@yale.edu). B. D. O. Anderson and C. Yu are with the Australian National University and National ICT Australia Ltd., Australia ({brian.anderson, brad.yu}@anu.edu.au). C. Yu is also with Shandong Computer Science Center, Jinan, China.

of algorithm development and convergence analysis. On the other hand, the deterministic approach forces one to consider worst case scenarios and has the potential of yielding algorithms which may outperform those obtained using the probabilistic approach. The aim of this paper is to present deterministic gossiping which do not utilize broadcasting and which generate sequences $x(0), x(1), x(2), \ldots$ which are guaranteed to converge exponentially fast to the limit vector which solves the distributed averaging problem.

II. Gossiping

The type of gossiping we want to consider involves a group of $n > 1$ agents labeled 1 to n. Each agent i has control over a real-valued scalar quantity x_i called a *gossip variable* which the agent is able to update. A *gossip* between agents i and j, written (i, j), occurs at time t if the values of both agents' variables at time $t + 1$ equal the average of their values at time t. In other words $x_i(t + 1) = x_j(t + 1) = \frac{1}{2}(x_i(t) + x_j(t))$. If agent i does not gossip at time t, its gossip variable does not change; thus in this case $x_i(t + 1) = x_i(t)$. Generally not every pair of agents is allowed to gossip. The edges of $\mathbb{A}$ specify which gossip pairs are allowable. In other words a gossip between agents i and j is *allowable* if (i, j) is an edge in $\mathbb{A}$. We sometimes call $\mathbb{A}$ an *allowable gossip graph*. Although in this paper we shall be interested primarily in gossiping protocols which stipulate that each agent is allowed to gossip with at most one of its neighbors at one time, as we shall see later, there is value in taking the time here to generalize the idea. Let us agree to call a subset $\mathcal{L}$ of $m > 1$ agent labels, a *neighborhood* if each pair of distinct labels in $\mathcal{L}$ are the labels of vertices in $\mathbb{A}$ which are connected. We say that the agents with labels in $\mathcal{L}$ perform a *gossip of order* m at time t if each updates its gossip variable to the average of all; that is, if $x_i(t + 1) = \frac{1}{m} \sum_{j \in \mathcal{L}} x_j(t),\ i \in \mathcal{L}$. A *generalized gossip* is a gossip of any order. A gossip without the modifier "generalized", will continue to mean a gossip of order 2.

One rule which sharply distinguishes a gossiping process from a more distributed averaging process is that in the case of gossiping, each agent is allowed to gossip with at most one of its neighbors at one time. This rule does not preclude the possibility of two or more pairs of agents gossiping at the same time, provided each of the two pairs have no agent in common. More precisely, two gossip pairs (i, j) and (k, m) are *noninteracting* if neither i nor j equals either k or m. When multiple noninteracting pairs of allowable gossips occur simultaneously, the simultaneous occurrence of all such gossips is called a *multi-gossip*. In other words a multi-gossip at time t is the set of all gossips which occur at time t with the understanding that each such pair is allowable and that any two such pairs are noninteracting. A *generalized multi-gossip* at time t is a finite set of generalized gossips with disjoint neighborhoods which occur simultaneously at time t.

A gossiping process can often be modeled by a discrete time linear system of the form $x(t + 1) = M(t)x(t),\ t = 0, 1, 2, \ldots$ where $x \in \mathbb{R}^n$ is a state vector of gossiping variables and $M(t)$ is a matrix characterizing how x changes as the result of the gossips which take place at time t; sometimes $M(t)$ depends on x although the notational dependence is often suppressed. If a single pair of distinct agents i and j gossip at time $t \geq 0$, then $M(t) = P_{ij}$ where P_{ij} is the $n \times n$ matrix for which $p_{ii} = p_{ij} = p_{ji} = p_{jj} = \frac{1}{2}$, $p_{kk} = 1, k \notin \{i, j\}$, and all remaining entries equal zero. We call such P_{ij} *single gossip primitive gossip matrices*. If at time t a multi-gossip occurs, then as a consequence of non-interaction, $M(t)$ is simply the product of the single gossip primitive gossip matrices corresponding to the individual gossips comprising the multi-gossip; moreover because of non-interaction, the primitive gossip matrices in the product commute with each other and so any given permutation of the primitive matrices in the product determines the same matrix P. We refer to P as the *primitive gossip matrix* determined by the multi-gossip under consideration.

The idea of a primitive gossip matrix extends naturally to generalized gossips. In particular, we associate with a neighborhood $\mathcal{L}$ the $n \times n$ matrix $P_{\mathcal{L}}$ where $p_{jk} = \frac{1}{m+1},\ j, k \in \mathcal{L}$, $p_{jj} = 1,\ j \notin \mathcal{L}$, and 0s elsewhere. We call $P_{\mathcal{L}}$ the *primitive gossip matrix* determined by $\mathcal{L}$. By the *graph induced by* $P_{\mathcal{L}}$, written $\mathbb{G}_{\mathcal{L}}$, we mean the spanning subgraph of $\mathbb{A}$ whose edge set is all edges in $\mathbb{A}$ which are incident on vertices with labels which are both in $\mathcal{L}$. More generally, if $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_k$ are k disjoint neighborhoods, the matrix $P_{\mathcal{L}_1} P_{\mathcal{L}_2} \cdots P_{\mathcal{L}_k}$ is the primitive gossip matrix determined by $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_k$ and the graph induced by $P_{\mathcal{L}_1} P_{\mathcal{L}_2} \cdots P_{\mathcal{L}_k}$ is the union of the induced graphs $\mathbb{G}_{\mathcal{L}_i},\ i \in \{1, 2, \ldots, k\}$. Note that the matrices in the product $P_{\mathcal{L}_1} P_{\mathcal{L}_2} \cdots P_{\mathcal{L}_k}$ commute because the $\mathcal{L}_i$ are disjoint so the order of the matrices in the product is not important for the definition to make sense. Note also that there are only finitely many primitive gossip matrices associated with $\mathbb{A}$.

A. Gossiping Sequences

Let $\gamma_1, \gamma_2, \ldots$ be an infinite sequence of multi-gossips corresponding to some or all of the edges in $\mathbb{A}$. Corresponding to such a sequence is a sequence of primitive gossip matrices $Q_1, Q_2, \ldots$ where Q_i is the primitive gossip matrices of the ith multi-gossip in the sequence. For given $x(0)$, such a gossiping matrix sequence *generates* the sequence of vectors

$$x(t) = Q_t Q_{t-1} \cdots Q_1 x(0), \quad t > 0 \tag{1}$$

which we call a *gossiping sequence*. We have purposely restricted this definition of a gossiping sequence to multi-gossip sequences, as opposed to *generalized* multi-gossip sequences, since we will only be dealing with algorithms involving multi-gossips. Our reason for considering generalized multi-gossips will become clear in a moment.

As will soon be obvious, the matrices Q_i in (1) are not necessarily the only primitive gossip matrices for which (1) holds. This non-uniqueness can play a crucial rule in understanding certain gossip protocols which are not linear iterations. To understand why this is so, let us agree to say that the transition $x(\tau) \longmapsto x(\tau+1)$ contains a *virtual gossip* if there is a neighborhood $\mathcal{L}$ for which $x_i(\tau) = x_j(\tau),\ i, j \in$

$\mathcal{L}$. We say that agent i has gossiped virtually with agent j at time t if i and j are both labels in $\mathcal{L}$. Thus while we are only interested in algorithms in which an agent may gossip with at most one neighbor at any one time, for such algorithms there may be times at which virtual gossips occur between an agent and one or more of its neighbors. Suppose that for some time $\tau < t$, the transition $x(\tau) \longmapsto x(\tau+1)$ contains such a virtual gossip and let $P_{\mathcal{L}}$ denote the primitive gossip matrix determined by $\mathcal{L}$. Then clearly $P_{\mathcal{L}}x(\tau) = x(\tau)$ which means that the matrix $Q_{\tau+1}$ in the product $Q_tQ_{t-1}\cdots Q_1$ can be replaced by the matrix $Q_{\tau+1}P_{\mathcal{L}}$ without changing the validity of (1). Moreover $Q_{\tau+1}P_{\mathcal{L}}$ will be a primitive gossip matrix if the neighborhoods which define $Q_{\tau+1}$ are disjoint with $\mathcal{L}$. The importance of this elementary observation is simply this. Without taking into account virtual gossips in equations such as (1), it may in some cases to be impossible to conclude that the matrix product $Q_tQ_{t-1}\cdots Q_1$ converges as $t \to \infty$ even though the gossip sequence $x(1), x(2), \ldots$ does. Later in this paper we will describe a gossip protocol for which this is true.

Prompted by the preceding, let us agree to say that a gossiping sequence satisfying (1) is *consistent* with a sequence of primitive gossip matrices $P_1, P_2, \ldots$ if

$$x(t) = P_tP_{t-1}\cdots P_1x(0), \qquad t > 0$$

It is obvious that if the sequence $x(t),\ t \geq 0$ is consistent with the sequence $P_1, P_2, \ldots$ and the latter converges, then so does the former. Given a gossip vector sequence, our task then is to find, if possible, a consistent, primitive gossip matrix sequence which is also convergent.

As we have already noted, $\mathbb{A}$ has associated with it a finite family of primitive gossip matrices and each primitive gossip matrix induces a spanning subgraph of $\mathbb{A}$. It follows that any finite sequence of primitive gossip matrix $P_1, P_2, \ldots, P_k$ *induces* a spanning subgraph of $\mathbb{A}$ whose edge set is the union of the edge sets of the graphs induced by all of the P_i. We say that the primitive gossip matrix sequence $P_1, P_2, \ldots, P_k$ is *complete* if the graph the sequence induces is a *connected* spanning subgraph of $\mathbb{A}$. An infinite sequence of primitive gossip matrices $P_1, P_2, \ldots$ is *repetitively complete* with period T, if each successive subsequence of length T in the sequence is complete. A gossiping sequence $x(t),\ t > 0$ is *repetitively complete* with period T, if there is a consistent sequence of primitive gossip matrices which is repetitively complete with period T. The importance of repetitive completeness is as follows.

Theorem 1: Suppose $P_1, P_2, \ldots$ is an infinite sequence of primitive gossip matrices which is repetitively complete with period T. There exists a real nonnegative number $\lambda < 1$, depending only on T and the P_t, for which

$$\lim_{t\to\infty} P_tP_{t-1}\cdots P_1x(0) = y_{\text{avg}}\mathbf{1}$$

as fast as λ^t converges to zero.

There are several different ways to prove this theorem using ideas from [2], [5], [6], [13], [14]. A proof of this theorem will be given in the full length version of this paper.

III. Request-Based Gossiping

Request-based gossiping is a gossiping process in which a gossip occurs between two agents whenever one of the two accepts a request to gossip placed by the other. The aim of this section is to discuss this process.

In a request-based gossiping process, a given agent i may gossip with one of its neighbors at time t only if t is either an "event time" of agent i or an "event time" of the neighbor which has made a request to gossip with agent i. By an *event time* of agent i is meant a time at which agent i may place a request to gossip with one of its neighbors. By an *event time interval* of agent i is meant the interval of time between two successive event times of agent i. For obvious reasons, we assume that the lengths of agent i's event time intervals are all bounded above by a finite positive number T_i. We write $\mathcal{T}_i$ for the set of event times of agent i and $\mathcal{T}$ for the union of the event time sequences of all n agents.

Conflicts leading to deadlocks can arise if an agent who has placed a request to gossip, at the same time receives a request to gossip from another agent. It is challenging to devise rules which resolve such conflicts while at the same time ensuring exponential convergence of the gossiping process. One way to avoid such conflicts is to assign event times off line so that no agent can receive a request to gossip at any of its own event times. There are several ways to do this which will be discussed below.

From time to time, agent i may have more than one neighbor to which it might be able to make a request to gossip with. Also from time to time, agent i may receive more than one request to gossip. While in such situations decisions about who to place requests with or whose request to accept can be randomized, in this paper we will examine only completely deterministic strategies. To do this we will assume that each agent i has ordered its neighbors in $\mathcal{N}_i$ according to some priorities so when a choice occurs between neighbors, agent i will always choose the one with highest priority.

Consider first the situation when the event times of each agent and each agent's neighbor priorities are chosen off line and are fixed throughout the gossiping process. Assume that the event times are chosen so that no agent can receive a request to gossip at any of its own event times. Our aim is to show that this arrangement can be problematic. The following protocol illustrates this.

Protocol I: At each event time $t \in \mathcal{T}$ the following rules apply for each $i \in \{1, 2, \ldots, n\}$:

1) If $t \in \mathcal{T}_i$, agent i places a request to gossip with that neighbor whose priority is the highest.
2) If $t \notin \mathcal{T}_i$, agent i does not place a request to gossip.
3) Each agent i receiving one or more requests to gossip must gossip with that requesting neighbor whose priority is the highest.
4) If $t \notin \mathcal{T}_i$ and agent i does not receive a request to gossip, it does not gossip.

The following example shows that this simple strategy will not necessarily lead to a consensus. Suppose that $\mathbb{A}$ is a path

graph with edges $(a,b),(b,c),(c,d)$. Assume that agents a and b have distinct event times and that agents a and c have the same event times as do agents b and d; note that this guarantees that no agent can receive a request to gossip at any of its own event times. To avoid ambiguities in decision making, suppose that agent b assigns a higher priority to a than to c and agent c assigns a higher priority to d than to b. Let t be an event time of agents a and c. Then at this time a places a request to gossip with b and c places a request to gossip with d. Since b and d receive no other requests, gossips take place between a and b and between c and d. Alternatively, if t is an event time of agents b and d, then at this time, b places a request to gossip with a and d places a request to gossip with c. Since a and c receive no other requests, gossips again take place between a and b and between c and d. Thus under no conditions is there ever a gossip between b and c, so the gossiping process will never reach a consensus. The reader may wish to verify that simply changing the priorities will not rectify this situation: For any choice of priorities, there will always be at least one gossip needed to reach a consensus, which will not take place.

The preceding example illustrates that fixed priorities can present problems. The global ordering proposed in [11] is one way to overcome them. In what follows we take an alternative approach.

In the light of Theorem 1 it is of interest to consider gossiping protocols which generate repetitively complete gossip sequences. Towards this end, let us agree to say that an agent i has completed a *round* of gossiping after it has gossiped with each neighbor in $\mathcal{N}_i$ at least once. Thus the finite sequence of primitive gossiping matrices corresponding to a finite sequence of multi-gossips for the entire group which has occurred over an interval of length T, will be complete if over the same period each agent in the group completes a round.

For the protocols which follow it will be necessary for each agent i to keep track of where it is in a particular round. To do this, agent i makes use of a recursively updated *neighbor queue* $\mathbf{q}_i(t)$ where $\mathbf{q}_i(\cdot)$ is a function from $\mathcal{T}$ to the set of all possible lists of the n_i labels in $\mathcal{N}_i$, the neighbor set of agent i. Roughly speaking, $\mathbf{q}_i(t)$ is a list of the labels of the neighbors of agent i which defines the queue of neighbors at time t which are in line to gossip with agent i. The updating of $\mathbf{q}_i(t)$ is straightforward: If neighbor j gossips with agent i at time t, the updated queue $\mathbf{q}_i(t+1)$ is obtained by moving agent j's label from its current position in $\mathbf{q}_i(t)$, to the end of the queue. If on the other hand, agent i does not gossip at time t, $\mathbf{q}_i(t+1)=\mathbf{q}_i(t)$.

A. Protocols

As noted earlier, it is helpful to have event time assignments which guarantee that no agent can receive a request to gossip at any of its own event times. One way to accomplish this is to use event time assignments which satisfy the following assumption.

Distinct neighbor event times assumption: For each $i \in \{1,2,\ldots,n\}$ and each $j \in \mathcal{N}_i$, $\mathcal{T}_i$ and $\mathcal{T}_j$ are disjoint sets.

Thus if this assumption holds, the event times of each agent are distinct from the event times of all of its neighbors. In all cases the largest number of distinct event time sequences which would need to be assigned to $\mathbb{A}$ to satisfy the distinct neighbor event times assumption is no greater than one plus the maximum vertex degree of $\mathbb{A}$ [15].

Under the distinct neighbor event times assumption, it is possible to ensure exponential convergence with the following protocol.

Protocol II: Suppose that the distinct neighbor event times assumption holds. At each event time $t \in \mathcal{T}$ the following rules apply for each $i \in \{1,2,\ldots,n\}$:

1) If $t \in \mathcal{T}_i$, agent i places a request to gossip with that neighbor whose label is at the front of the queue $\mathbf{q}_i(t)$.
2) If $t \notin \mathcal{T}_i$, agent i does not place a request to gossip.
3) Each agent i receiving one or more requests to gossip must gossip with that requesting neighbor whose label is closest to the front of the queue $\mathbf{q}_i(t)$.
4) If $t \notin \mathcal{T}_i$ and agent i does not receive a request to gossip, it does not gossip.

Proposition 1: Let $\mathcal{E}$ denote the set of all edges (i,j) in $\mathbb{A}$. Suppose that the distinct neighbor event times assumption holds and that all agents in the group adhere to Protocol II. Then the infinite gossiping sequence generated will be repetitively complete with period

$$T = \max_{(i,k)\in\mathcal{E}} \min \left\{ T_i \sum_{j\in\mathcal{N}_i} n_j,\ T_k \sum_{j\in\mathcal{N}_k} n_j \right\}$$

The proof of Proposition 1 can be found in [15].

A disadvantage of Protocol II is that it requires the distinct neighbor event times assumption. This assumption can only be satisfied by off-line assignment of event times for each agent, and in some applications such an off-line assignment may be undesirable. In a recent doctoral thesis [16], a clever gossiping protocol is proposed which does not require the distinct neighbor event times assumption. The protocol avoids deadlocks and achieves consensus exponentially fast. A disadvantage of the protocol in [16] is that it requires each agent to obtain the values of *all* of its neighbors' gossip variables at each clock time. By exploiting one of the key ideas in [16] together with the notion of an agent's neighbor queue $\mathbf{q}_i(t)$ defined earlier, it is possible to obtain a gossiping protocol which also avoids deadlocks and achieves consensus exponentially fast but without requiring each agent to obtain the values of all of its neighbors' gossip variables at each iteration.

In the sequel we will outline a gossiping algorithm in which at time t, each agent i has a single *preferred neighbor* whose label $i^*(t)$ is in the front of queue $\mathbf{q}_i(t)$. At time t each agent i transmits to its preferred neighbor its label i and the current value of its gossip variable $x_i(t)$. Agent i then transmits the current value of its gossip variable to those agents which have agent i as their preferred neighbor; these neighbors plus neighbor $i^*(t)$ are agent i's *receivers* at time t. They are the neighbors of agent i who know

the current gossip value of agent i. Agent i is presumed to have placed a request to gossip with its preferred neighbor $i^*(t)$ if $x_i(t) > x_{i^*(t)}(t)$; agent i is a *requester* of agent $i^*(t)$ whenever this is so. Note that while an agent has exactly one preferred neighbor, it may at the same time have anywhere from zero to n_i requesters, where n_i is the number of neighbors of agent i.

Protocol III: Between clock times t and $t+1$ each agent i performs the steps enumerated below in the order indicated. Although the agents' actions need not be precisely synchronized, it is understood that for each $k \in \{1,2,3\}$ all agents complete step k before any embark on step $k+1$.

1) **1st Transmission:** Agent i sends its label i and its gossip value $x_i(t)$ to its current preferred neighbor. At the same time agent i receives the labels and corresponding gossip values from all of those neighbors which have agent i as their current preferred neighbor.
2) **2nd Transmission:** Agent i sends its current gossip value $x_i(t)$ to those neighbors which have agent i as their current preferred neighbor.
3) **Acceptances:**
   a) If agent i has not placed a request to gossip but has received at least one request to gossip, then agent i sends an acceptance to that particular requesting neighbor whose label is closest to the front of the queue $\mathbf{q}_i(t)$.
   b) If agent i either has placed a request to gossip or has not received any request to gossip, then agent i does not send out an acceptance.
4) **Gossip variable and queue updates:**
   a) If agent i either sends an acceptance to or receives an acceptance from neighbor j, then agent i gossips with neighbor j by setting

$$x_i(t+1) = \frac{x_i(t) + x_j(t)}{2}$$

   Agent i updates its queue by moving j and the labels of all of its current receivers k, if any, for which $x_k(t) = x_i(t)$ from their current positions in $\mathbf{q}_i(t)$ to the end of the queue while maintaining their relative order.
   b) If agent i has not sent out an acceptance nor received one, then agent i does not update the value of $x_i(t)$. In addition, $\mathbf{q}_i(t)$ is not updated except when agent i's gossip value equals that of at least one of its current receivers. In this special case agent i moves the labels of all of its current receivers k for which $x_k(t) = x_i(t)$ from their current positions in $\mathbf{q}_i(t)$ to the end of the queue, while maintaining their relative order.

In summary,

- For agent i to place a request to gossip, the current value of its gossip variable must be larger than that of its current preferred neighbor.
- For a gossip to occur between two agents i and j at time t, one – say i – must be the current preferred neighbor of the other {i.e., $i = j^*(t)$}, $x_j(t)$ must be larger than $x_i(t)$, and j must be the label of the neighbor of agent i with highest priority which is placing a request to gossip with agent i.
- For agent i to update its queue, it must either gossip with a neighbor j or, if not, it's current gossip value must equal that of at least one of its receivers.

Transmissions required: During step 1, each agent sends a transmission to its preferred neighbor so the total number of transmissions required for all n agents to complete step 1 is n. During step 2, each neighbor of agent i which has agent i as its current preferred neighbor sends a transmission to agent i so the total of transmissions required for all n agents to complete step 2 is also n. The total number of transmissions of all agents required to complete step 3a is clearly no greater than $\frac{n}{2}$. Thus the total number of transmissions per iteration to carry out the protocol just described is no greater than $\frac{5}{2}n$. With a broadcasting protocol such as the one considered in [16] the total number of transmissions per iteration is nd_{avg} where d_{avg} is the average vertex degree of the underlying graph $\mathbb{A}$. Thus for allowable gossip graphs with average vertex degree exceeding $\frac{5}{2}$, fewer transmissions are required per iteration to do averaging with the protocol under consideration than are required per iteration to do averaging via broadcasting.

Theorem 2: Every sequence of gossip vectors $x(t),\ t > 0$ generated by protocol III is repetitively complete with period no greater than the number of edges of $\mathbb{A}$.

Theorems 1 and 2 thus imply that every sequence of gossip vectors generated by protocol III converges to the desired limit point exponentially fast at a rate no worse that some finite number $\lambda < 1$ which depends only $\mathbb{A}$. Calculation of this worst case bound is a subject for future research.

To prove Theorem 2, we need a few ideas. First note that step 4 of the protocol stipulates that agent i must update its queue whenever its current gossip value equals that of one of its neighbors. We say that agent i *gossips virtually* with neighbor j at time t if the current gossip values of both agents are the same. Note that while an agent can gossip with at most one agent at time t, it can gossip virtually with as many as n_i at the same time. To proceed, we need to generalize slightly the idea of a round. We say that an agent i has completed a round of gossiping after it has gossiped or virtually gossiped with each neighbor in $\mathcal{N}_i$ at least once. Thus the finite sequence of primitive gossiping matrices corresponding to a finite sequence of multi-gossips and virtual multi-gossips for the entire group which has occurred over an interval of length T, will be complete if over the same period each agent in the group completes a round. Thus Theorem 2 will be true if every agent completes a round in a number of iterations no larger than the number of edges of $\mathbb{A}$. The following proposition asserts that this is in fact the case.

Proposition 2: Let m be the number of edges in $\mathbb{A}$. Then within m iterations every agent will have gossiped or virtually gossiped at least once with each of its neighbors.

To prove this proposition we will make use of the follow-

ing two lemmas.

Lemma 1: Suppose that all n agents follow protocol III. Then at each time t, at least one gossip or virtual gossip must occur.

Lemma 2: Let t be fixed and suppose that $\mathbb{G}$ is a spanning subgraph of $\mathbb{A}$ with at least one edge. For each $i \in \{1, 2, \ldots, n\}$ write $\mathcal{N}_i$ for the set of labels of the vertices adjacent to vertex i in $\mathbb{A}$ and $\mathcal{M}_i$ for the set of labels of the vertices adjacent to vertex i in $\mathbb{G}$. Let $\mathcal{N}_i - \mathcal{M}_i$ denote the complement of $\mathcal{M}_i$ in $\mathcal{N}_i$. Suppose that for each $i \in \{1, 2, \ldots, n\}$, each label in $\mathcal{M}_i$, if any, is closer to the front of $\mathbf{q}_i(t)$ than all the labels in $\mathcal{N}_i - \mathcal{M}_i$. Then there must be an edge (i, j) within $\mathbb{G}$ such that at time t, neighboring agents i and j either gossip or gossip virtually.

The proofs of Lemma 1 and Lemma 2 are omitted due to space limitations; they will be given in a full length version of this paper.

Proof of Proposition 2: If $m = 1$, there can be only two agents so $n = 2$. In view of Lemma 1, Proposition 2 must clearly be true for this case.

Suppose $m > 1$. Fix t and let $\mathcal{E}_0$ be the edge set of $\mathbb{A}$. For $k \in \{1, 2, \ldots, m\}$ let $\mathcal{E}_k$ denote the set of all edges (i, j) in $\mathcal{E}_0$ for which agents i and j have gossiped or virtually gossiped at least once within k iterations starting at time t. Fix k. If $\mathcal{E}_k = \mathcal{E}_0$, then each agent will have gossiped or virtually gossiped at least once with each of its neighbors within k iterations. Since $k \leq m$, each agent will have gossiped or virtually gossiped at least once with each of its neighbors within m iterations starting at time t.

Now suppose that $\mathcal{E}_k \neq \mathcal{E}_0$ in which case the complement of $\mathcal{E}_k$ in $\mathcal{E}_0$, namely $\mathcal{E}_0 - \mathcal{E}_k$, is nonempty. Let $\mathbb{G}_k$ denote the spanning subgraph of $\mathbb{A}$ with edge set $\mathcal{E}_0 - \mathcal{E}_k$. For each $i \in \{1, 2, \ldots, n\}$ write $\mathcal{N}_i$ for the set of labels of the vertices adjacent to vertex i in $\mathbb{A}$ and $\mathcal{M}_i$ for the set of labels of the vertices adjacent to vertex i in $\mathbb{G}_k$. Let $\mathcal{N}_i - \mathcal{M}_i$ denote the complement of $\mathcal{M}_i$ in $\mathcal{N}_i$. For each label $j \in \mathcal{N}_i - \mathcal{M}_i$, if any, $(i, j) \in \mathcal{E}_k$ which means that agents i and j have gossiped or virtually gossiped at least once within k iterations starting at time t. On the other hand, if there is vertex $j \in \mathcal{M}_i$, then this vertex labels an agent which has not gossiped or virtually gossiped with agent i within k iterations. Protocol III stipulates that each label in $\mathcal{M}_i$ is closer to the front of $\mathbf{q}_i(t+k)$ than all the labels in $\mathcal{N}_i - \mathcal{M}_i$. Since i is arbitrary, this is true for all $i \in \{1, 2, \ldots, n\}$. It follows from Lemma 2 that there is an edge (a, b) in $\mathcal{E}_0 - \mathcal{E}_k$ such that at time $t + k$, neighboring agents a and b either gossip or virtually gossip.

By hypothesis $\mathcal{E}_k \neq \mathcal{E}_0$. Since $\mathcal{E}_k \supset \mathcal{E}_j$ for $j \in \{1, 2, \ldots, k-1\}$, it must be true that $\mathcal{E}_j \neq \mathcal{E}_0$ for $j \in \{1, 2, \ldots, k\}$. Thus the preceding argument applies for all $j \in \{1, 2, \ldots, k\}$, so for each such j there must be an edge (a_j, b_j) in $\mathcal{E}_0 - \mathcal{E}_j$ such that at time $t + j$, neighboring agents a_j and b_j either gossip or virtually gossip. Clearly $(a_j, b_j) \notin \{(a_1, b_1), (a_2, b_2), \ldots, (a_{j-1}, b_{j-1})\}$ for $j \in \{2, 3, \ldots, k\}$ because $\{(a_1, b_1), (a_2, b_2), \ldots, (a_{j-1}, b_{j-1})\} \subset \mathcal{E}_j$ and $(a_j, b_j) \in \mathcal{E}_0 - \mathcal{E}_j$. It follows that $(a_1, b_1), (a_2, b_2), \ldots, (a_k, b_k)$ are distinct edges in $\mathbb{A}$.

The preceding argument implies that at the end of k iterations, either $\mathcal{E}_k = \mathcal{E}_0$ or k distinct gossips/virtual gossips have taken place. If the former is true, then each agent will have gossiped or virtually gossiped at least once with each of its neighbors within k and therefore m iterations. If the later is true, then k must be less than m and k distinct gossips/virtual gossips will have taken place within k iterations. Clearly this process can be continued until for some integer $k \leq m$, $\mathcal{E}_k = \mathcal{E}_0$ in which case each agent will have gossiped or virtually gossiped at least once with each of its neighbors within k and therefore m iterations. ■

IV. Concluding Remarks

One of the problems with the idea of gossiping, which apparently is not widely appreciated, is that it is difficult to devise provably correct gossiping protocols which are guaranteed to avoid deadlocks without making restrictive assumptions. The research in this paper and in [11] and [16] contributes to our understanding of this issue and how to deal with it.

References

[1] J. N. Tsitsiklis. *Problems in decentralized decision making and computation*. PhD thesis, Department of Electrical Engineering and Computer Science, MIT, 1984.

[2] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003.

[3] R. Olfati-Saber and R. M. Murray. Consensus seeking in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.

[4] V. D. Blondel, J. M. Hendrichkx, A. Olshevsky, and J. N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 2996–3000, 2005.

[5] L. Moreau. Stability of multi-agent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50(2):169–182, 2005.

[6] W. Ren and R. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, 50(5):655–661, 2005.

[7] L. Xiao and S. Boyd. Fast linear iterations for distributed averaging. *Systems and Control Letters*, 53(1):65–78, 2004.

[8] A. Kashyap, T. Basar, and R. Srikant. Quantized consensus. *Automatica*, 43(7):1192–1203, 2007.

[9] L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *Proceedings of the 4th International Conference on Information Processing in Sensor Networks*, pages 63–70, 2005.

[10] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

[11] M. Mehyar, D. Spanos, J. Pongsajapan, S. H. Low, and R. M. Murray. Asynchronous distributed averaging on communication networks. *IEEE/ACM Transactions on Networking*, 15(3):512–520, 2007.

[12] B. D. O. Anderson, C. Yu, and A. S. Morse. Convergence of periodic gossiping algorithms. In *Perspectives in Mathematical System Theory, Control, and Signal Processing*, pages 127–138. Springer, 2010.

[13] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis. On distributed averaging algorithms and quantization effects. *IEEE Transactions on Automatic Control*, 54(11):2506–2517, 2009.

[14] J. Liu, A. S. Morse, B. D. O. Anderson, and C. Yu. Contractions for consensus processes. In *Proceedings of the 50th IEEE Conference on Decision and Control*, 2011. to appear.

[15] J. Liu, S. Mou, A. S. Morse, B. D. O. Anderson, and C. Yu. Deterministic gossiping. *Proceedings of the IEEE*, 2011. to appear.

[16] A. Olshevsky. *Efficient Information Aggregation Strategies for Distributed Control and Signal Processing*. PhD thesis, Department of Electrical Engineering and Computer Science, MIT, 2010.

INVITED PAPER

Deterministic Gossiping

Gossip algorithms can provide information exchange and computation for autonomous vehicles in a group, where each vehicle must make estimates and decisions, while ensuring consensus at the group level.

By Ji Liu, *Student Member IEEE*, Shaoshuai Mou, *Student Member IEEE*, A. Stephen Morse, *Life Fellow IEEE*, Brian D. O. Anderson, *Life Fellow IEEE*, and Changbin Yu, *Senior Member IEEE*

ABSTRACT | For the purposes of this paper, "gossiping" is a distributed process whose purpose is to enable the members of a group of autonomous agents to asymptotically determine, in a decentralized manner, the average of the initial values of their scalar gossip variables. This paper discusses several different deterministic protocols for gossiping which avoid deadlocks and achieve consensus under different assumptions. First considered is *T*-periodic gossiping which is a gossiping protocol which stipulates that each agent must gossip with the same neighbor exactly once every *T* time units. Among the results discussed is the fact that if the underlying graph characterizing neighbor relations is a tree, convergence is exponential at a worst case rate which is the same for all possible *T*-periodic gossip sequences associated with the graph. Many gossiping protocols are request based which means simply that a gossip between two agents will occur whenever one of the two agents accepts a request to gossip placed by the other. Three deterministic request-based protocols are discussed. Each is guaranteed to not deadlock and to always generate sequences of gossip vectors which converge exponentially fast. It is shown that worst case convergence rates can be characterized in terms of the second largest *singular* values of suitably defined doubly stochastic matrices.

KEYWORDS | Consensus; distributed averaging; nonhomogeneous Markov chains; stochastic matrices

I. INTRODUCTION

There has been considerable interest recently in developing algorithms for distributing information among the members of a group of sensors or mobile autonomous agents via local interactions. Notable among these are those algorithms intended to cause such a group to reach a consensus in a distributed manner [1]–[7]. One particular type of consensus process which has received much attention lately is called distributed averaging [8]. In its simplest form, distributed averaging deals with a network of $n > 1$ *agents* and the constraint that each agent i is able to communicate only with certain other agents called agent i's *neighbors*. Neighbor relations are described by a simple, connected graph $\mathbb{N}$ in which vertices correspond to agents and edges indicate neighbor relations. Thus, the neighbors of an agent i have the same labels as the vertices in $\mathbb{N}$ which are adjacent to vertex i. Initially, each agent i has or acquires a real number y_i which might be a measured temperature or something similar. The *distributed averaging problem* is to devise a protocol which will enable each agent to compute the average $y_{\text{avg}} = (1/n)\sum_{i=1}^{n} y_i$ using only information acquired from its neighbors. There are many variants of this problem. For example, instead of real numbers, the y_i may be integer-valued [9]. Another variant assumes that the edges of $\mathbb{N}$ change over time [10]. This paper considers the case when the y_i are real numbers and $\mathbb{N}$ does not depend on time.

As noted in [8], the distributed averaging problem can be solved, in principle, by "flooding"; that is, by propagating across the network over time the values of all of the y_i. Armed with knowledge of all of these values, each agent

Manuscript received May 20, 2010; revised December 6, 2010; accepted May 25, 2011. Date of publication August 4, 2011; date of current version August 19, 2011. The work of L. Jiu, S. Mou, and A. S. Morse was supported by the U.S. Army Research Office, the U.S. Air Force Office of Scientific Research, and the National Science Foundation. The work of B. D. O. Anderson was supported by the Australian Research Council's Discovery Project DP-110100538 and the National ICT Australia—NICTA. NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. The work of C. B. Yu was supported by the Australian Research Council through a Queen Elizabeth II Fellowship and DP-110100538 and by the Overseas Expert Program of Shandong Province, China. The work of B. D. O. Anderson and C. B. Yu was also supported by the U.S. Air Force Research laboratory Grant FA2386-10-1-4102.

J. Liu, **S. Mou**, and **A. S. Morse** are with Yale University, New Haven, CT 06520 USA (e-mail: ji.liu@yale.edu; shaoshuaimou@yale.edu; as.morse@yale.edu).

B. D. O. Anderson is with the Australian National University, Canberra, A.C.T. 0200, Australia and also with the National ICT Australia (NICTA), Canberra Research Laboratory, Canberra, A.C.T. 2601, Australia.

C. Yu is with the Australian National University, Canberra, A.C.T. 0200, Australia, the National ICT Australia (NICTA), Canberra Research Laboratory, Canberra, A.C.T. 2601, Australia, and also Shandong Computer Center, Jinan, China.

Digital Object Identifier: 10.1109/JPROC.2011.2159689

is thus able to compute y_{avg}. A more sophisticated approach to the problem is for each agent to use a linear iterative update rule of the general form

$$x_i(t+1) = w_{ii}x_i(t) + \sum_{j\in\mathcal{N}_i} w_{ij}x_j(t), \qquad x_i(1) = y_i$$

where t is a discrete time index, $x_i(t)$ is agent i's current estimate of y_{avg}, the w_{ij} are real-valued weights, and $\mathcal{N}_i$ is the set of labels of the neighbors of agent i. In [8], several methods are proposed for choosing the w_{ij}. One particular choice, which defines what has come to be known as the Metropolis algorithm, requires only local information to define the w_{ij}. Algorithms of this type, which require each agent to communicate with all of its neighbors on each iteration, are sometimes called *broadcast* algorithms.

An alternative approach to distributed averaging, which typically does not involve broadcasting, exploits a form of "gossiping" [11] specifically tailored to the distributed averaging problem. The idea of gossiping is very simple. A pair of neighbors with labels i and j are said to *gossip* at time t if both $x_i(t+1)$ and $x_j(t+1)$ are set equal to the average of $x_i(t)$ and $x_j(t)$. Each agent is allowed to gossip with at most one neighbor at one time. Under appropriate assumptions, algorithms which possess this simple property can be shown to solve the distributed averaging problem. Gossiping algorithms do not necessarily involve broadcasting and thus have the potential to require less transmissions per iteration than broadcast algorithms. Of course, one would not expect gossip algorithms to converge as fast as broadcast algorithms.

Gossiping might find application in many contexts. For example, suppose that a spatially distributed network of temperature sensors has been deployed in such a way so that each sensor can communicate with nearby sensors (i.e., neighbors). Suppose that at some specific time t_1 all sensors take readings and use these readings as initial estimates of the average temperature across the network. At subsequent clock times, each sensor then passes its current estimate to one of its neighbors who in turn uses this estimate to update its own estimate with the goal of ultimately arriving at the average value of the temperature across the network. Gossiping is a process for recursively carrying out these computations.

Implementation of any gossiping protocol necessarily involves some degree of centralization. In particular, for the aforementioned sensor network averaging task to make sense each sensor must be instructed by a centralized manager to take a temperature reading at the same time $t = t_1$ as the rest. In some cases, it may be useful to take centralization further. For example, to facilitate gossiping it may be helpful in some instances to centrally define a network-wide sensor ordering by assigning offline to each sensor a unique priority number with the understanding that each sensor can make use of the priority numbers of its neighbors to carry out its role in the gossiping process [12]. Another idea requiring some degree of centralization is to assign offline to each agent, a specific sequence of times at which the agent may gossip (Section IV-A). One might carry centralization one step further by specifying offline for each agent which neighbor the agent is to gossip with at each clock time (Section III). This of course could be done only in a network whose population of sensors does not change over time in a predictable manner. For networks whose members change with time, one might consider other ideas and assumptions. For example, for the temperature sensing network, one might try implementing a gossiping protocol which assumes that at each clock time each sensor can acquire the current temperature estimates of *all* of its neighbors [13]. A refinement of this protocol which requires each agent to acquire the current temperature of only a subset of its neighbors at each clock time is discussed in Section IV-A. In the end, which assumptions make sense depends on the specific application.

Although gossiping is a form of consensus, it differs from consensus in several important ways. First, the goal of consensus is to agree on the value of some quantity whereas the goal of gossiping is to compute the average of the initial values of the x_i, henceforth called *gossip variables*. Second, unlike consensus processes, gossiping processes are invariably designed so that the sum total of all gossip variables remains constant from clock time to clock time. This has a simple but important consequence: If the sum total of all gossip variables remains constant and a consensus is reached in that all gossip variables converge to the same value, then this value *must* be the average of the initial values of all gossip variables in the network. In gossiping, the sum total of all gossip variables is kept constant by requiring agents to always gossip in pairs using averaging. Although this simple idea keeps constant the sum of gossip variables across the network, the idea comes with a price in that a deadlock may well occur unless specific precautions are built into the protocol to preclude this.

The specific sequence of gossips which occurs during a given gossiping process might be determined either probabilistically [11], [14] or deterministically [12], [15], depending on the problem of interest. Deterministic gossiping protocols are intended to guarantee that under all conditions, a consensus will be achieved asymptotically whereas probabilistic protocols aim at achieving consensus asymptotically with probability one. Both approaches have merit. The probabilistic approach is typically somewhat easier both in terms of algorithm development and convergence analysis. On the other hand, the deterministic approach forces one to consider worst case scenarios and has the potential of yielding algorithms which may outperform those obtained using the probabilistic approach. This paper takes the deterministic approach.

Of particular interest is the rate at which a sequence of agent gossip variables converge to a common value. The convergence rate question for more general deterministic

consensus problems has been studied in [16] and [17]. In [11] and [14], the convergence rate question is addressed for gossiping algorithms in which the sequence of gossip pairs under consideration is determined probabilistically. A modified gossiping algorithm intended to speed up convergence is proposed in [18] without proof of correctness, but with convincing experimental results. The algorithm has recently been analyzed in [19]. Recent results concerning convergence rates appear in [15], [20], and [21] for periodic gossiping and in [22]–[24] for deterministic aperiodic gossiping. This paper presents a more comprehensive treatment of the work in [15] and [24].

A typical gossiping process can usually be modeled as a discrete time linear system of the form $x(t+1) = M(t)x(t)$, $t = 1, 2, \ldots$, where x is a vector of agent gossip variables x_i and each value of $M(t)$ is a specially structured doubly stochastic matrix (Section II). Roughly speaking, a finite sequence of gossip pairs is "complete" if the corresponding set of edges in $\mathbb{N}$ forms a connected spanning subgraph. A complete gossip sequence is minimally complete if there is no other complete gossip sequence of shorter length (Section II-B). An infinite sequence of gossips is "repetitively complete" with period T if each successive subsequence of gossips of length T in the sequence is complete. The gossip variable sequences associated with repetitively complete gossip sequences converge exponentially fast (Section II-C). Repetitively complete gossip sequences which are also periodic with period T are treated in Section III. The worst case convergence rate of any such sequence is determined by T and by the second largest eigenvalue (in magnitude) of the stochastic matrix which the gossips define over a period. In the case when $\mathbb{N}$ is a tree and the sequence of gossips over a period is minimally complete, the value of this eigenvalue does not depend on the order in which gossips over a period take place. A proof of this surprising result is given in [25].

Most gossiping protocols are "request-based." By *request-based* gossiping is meant a gossiping process in which a gossip occurs between two agents whenever one of the two accepts a request to gossip placed by the other (Section IV). An agent's "event times" are the times at which the agent makes requests to gossip. In Section IV-A, a request-based protocol is given which generates repetitively complete (and thus exponentially convergent) gossip sequences under the assumption that the event times of each agent are different than the event times of all of its neighbors. A more refined repetitively complete gossiping protocol not requiring this assumption is discussed in Section IV-A. The protocol is inspired by ideas put forth in [13].

It is shown in Section IV-C that the worst case convergence rate of a repetitively complete gossip sequence with period T can be characterized in terms of a suitably defined seminorm of the stochastic matrix S determined by the subsequence of gossips occurring over a given period. A specific goal of this paper is to find a seminorm with respect to which S is a contraction. The role played by seminorms in characterizing convergence rate is explained in Section IV-C. Three different types of seminorms are considered in Section IV-C. Each is compared to the well-known coefficient of ergodicity which plays a central role in the study of convergence rates for nonhomogeneous Markov chains [26]. Somewhat surprisingly, it turns out that a particular Euclidean seminorm on $\mathbb{R}^{n\times n}$ has the required property—namely that in this seminorm, the stochastic matrix S determined by any complete gossip sequence is a contraction (Section IV-D). This particular seminorm turns out to be the second largest singular value of S.

II. GOSSIPING

Gossiping, as used here, is a form of distributed computation whose purpose is to calculate the average value of a set of numbers or measurements existing at the "nodes" of a distributed network. The type of gossiping we want to consider involves a group of $n > 1$ agents labeled 1 to n. Each agent i has control over a real-valued scalar quantity x_i called a *gossip variable* which the agent is able to update from time to time. A *gossip* between agents i and j, written (i, j), occurs at time $t \in \{1, 2, \ldots\}$ if the values of both agents' variables at time $t+1$ equal the average of their values at time t. In other words $x_i(t+1) = x_j(t+1) = (1/2)(x_i(t) + x_j(t))$. If agent i does not gossip at time t, its gossip variable does not change; thus in this case $x_i(t+1) = x_i(t)$. Agents may only gossip with their neighbors. Agent j is a *neighbor* of agent i if (i, j) is an edge in a given simple, undirected, n-vertex graph $\mathbb{N}$ called a *neighbor graph*; we use the symbol $\mathcal{N}_i$ to denote the set of labels of the neighbors of agent i. One rule which sharply distinguishes a gossiping process from a more general consensus process is that in the case of gossiping, each agent is allowed to gossip with at most one of its neighbors at one time. This rule does not preclude the possibility of two or more pairs of agents gossiping at the same time, provided each of the two pairs have no agent in common. More precisely, two gossip pairs (i, j) and (k, m) are *noninteracting* if neither i nor j equals either k or m. When multiple noninteracting pairs of gossips occur simultaneously, the simultaneous occurrence of all such gossips is called a *multigossip*. In other words, a multigossip at time t is the set of all gossips which occur at time t with the understanding that any two such pairs are noninteracting. A central goal of gossiping is for the n agents to reach a consensus in the sense that all n gossip variables ultimately reach the same value in the limit as $t \to \infty$. For this to be possible, no matter what the initial values of the gossip variables are, it is clearly necessary that $\mathbb{N}$ be a connected graph. We assume that this is so.

A gossiping process can be modeled by a discrete time linear system of the form $x(t+1) = M(t)x(t)$, $t = 1, 2, \ldots$, where $x \in \mathbb{R}^n$ is a state vector of gossip variables and $M(t)$

is a matrix characterizing how x changes as the result of the gossips which take place at time t. If a single pair of distinct agents i and j gossip at time $t \geq 1$, then $M(t) = P_{ij}$ where P_{ij} is the $n \times n$ matrix for which $p_{ii} = p_{ij} = p_{ji} = p_{jj} = (1/2)$, $p_{kk} = 1$, $k \notin \{i, j\}$ and all remaining entries equal zero. We call such P_{ij} *single-gossip matrices*. It will be convenient to include in the set of single-gossip matrices, the $n \times n$ identity matrix; the identity matrix can be thought of as the correct update matrix to model the situation when no gossips occur at time t. If at time t a multigossip occurs, then as a consequence of noninteraction, $M(t)$ is simply the product of the single-gossip matrices corresponding to the individual gossips comprising the multigossip; moreover because of noninteraction, the single-gossip matrices in the product commute with each other and so any given permutation of the single-gossip matrices in the product determines the same matrix P. We refer to P as the *gossip matrix* determined by the multigossip under consideration. Thus, for example, if at time t agents a, b, and c gossip with agents d, e, and f, respectively, then the gossip matrix determined is $P = P_{ad}P_{be}P_{cf}$. Thus, in this case, $M(t) = P$ and $x(t+1) = M(t)x(t)$.

A. Doubly Stochastic Matrices

Each single-gossip matrix is a nonnegative matrix whose row sums and column sums all equal one. Matrices with these two properties are called *doubly stochastic*. Note that the type of doubly stochastic matrices which characterize single gossips (i.e., single-gossip matrices) has two additional properties—it is symmetric and its diagonal entries are all positive. The same is true for the type of doubly stochastic matrices which characterize multigossips. Doubly stochastic matrices are special types of "stochastic matrices" where by a *stochastic matrix* is meant a nonnegative $n \times n$ matrix whose row sums all equal one. It is easy to see that a nonnegative matrix S is stochastic if and only if $S\mathbf{1} = \mathbf{1}$ where $\mathbf{1} \in \mathbb{R}^n$ is a column vector whose entries are all ones. Similarly, a nonnegative matrix S is doubly stochastic if and only if $S\mathbf{1} = \mathbf{1}$ and $S'\mathbf{1} = \mathbf{1}$. Using these characterizations it is easy to prove that the class of stochastic matrices in $\mathbb{R}^{n\times n}$ is compact and closed under multiplication as is the class of doubly stochastic matrices in $\mathbb{R}^{n\times n}$. It is also true that the class of nonnegative matrices in $\mathbb{R}^{n\times n}$ with positive diagonal entries is closed under multiplication.

Mathematically, reaching a consensus by means of an infinite sequence of gossips or multigossips modeled by a corresponding infinite sequence of gossip matrices $M(1), M(2), M(3), \ldots$ means that the sequence of matrix products $M(1), M(2)M(1), M(3)M(2)M(1), \ldots$ converges to a matrix of the form $\mathbf{1}c$. It turns out that if convergence occurs, the limit matrix $\mathbf{1}c$ is also a doubly stochastic matrix; this means that $c = (1/n)\mathbf{1}'$ and consequently that all n gossip variables will have converged to the average of their initial values. This particular fact further distinguishes a gossiping process from a more general consensus process, since in a consensus process the value to which all consensus variables typically converge is not necessarily the average of their initial values.

B. Gossiping Sequences

By a *gossiping sequence* is meant a sequence of individual gossips corresponding to some or all of the edges in a given neighbor graph $\mathbb{N}$. Corresponding to any given sequence of gossips $(i_1, j_1), (i_2, j_2), \ldots$ is a sequence of single-gossip matrices $P_{i_1j_1}, P_{i_2j_2}, \ldots$ whose product $\cdots P_{i_2j_2}P_{i_1j_1}$ defines the mapping which assigns to any given initial vector of gossip variables, the vector of gossip variables which results from the gossips in the sequence. We call any such matrix product a *gossip matrix*. It is thus clear that a given neighbor graph has associated with it a family of gossip matrices whose members are all products of all combinations of single-gossip matrices of all lengths. These are the gossip matrices determined by $\mathbb{N}$. Conversely, any given sequence of individual gossips (or corresponding product of single-gossip matrices) *induces* a spanning subgraph of $\mathbb{N}$ whose edges correspond to the gossips in the sequence. We say that a gossip sequence or corresponding gossip matrix is *complete* if the graph the gossips in the sequence induce is a connected spanning subgraph within $\mathbb{N}$. A gossip sequence and corresponding gossip matrix is *minimally complete*, if it is complete and if there is no other complete gossip sequence of shorter length. It is easy to see that a nonredundant[1] gossip sequence is minimally complete if and only if the subgraph of $\mathbb{N}$ that it induces is a minimal spanning tree in $\mathbb{N}$. In the special but important case when $\mathbb{N}$ is itself a tree $\mathbb{T}$, more can be said. In this case, a minimally complete gossip sequence is one in which, for each edge in $\mathbb{T}$, there is exactly one corresponding individual gossip in the sequence.

The preceding ideas extend in a natural way to sequences of multigossips. Corresponding to any given sequence of multigossips $\gamma_1, \gamma_2, \ldots$ is a sequence of gossip matrices $Q_1, Q_2, \ldots$ where Q_i is the product of the single-gossip matrices of the individual gossips in the ith multigossip in the sequence. The product $\cdots Q_2Q_1$ thus defines the mapping which assigns to any given initial vector of gossip variables the vector of gossip variables which results from the multigossips in the sequence. Clearly any such matrix product is also a product of single-gossip matrices and thus is a bona fide gossip matrix.

Extending the concept of nonredundancy, we say that a multigossip sequence is *nonredundant* if no individual gossip occurs in more than one multigossip in the sequence. Nonredundant multigossip sequences are clearly of finite length in that the length of each is no larger than the number of edges of $\mathbb{N}$. The *graph induced by* a multigossip sequence Σ, written $\mathbb{N}_\Sigma$, is the spanning subgraph of $\mathbb{N}$ whose edges correspond to all of the gossips in all of the multigossips in the sequence. A multigossip sequence

[1]A gossip sequence is *nonredundant* if each gossip in the sequence occurs at most once.

is *complete* if the graph $\mathbb{N}_\Sigma$ which it induces is a connected spanning subgraph of $\mathbb{N}$. Σ is *minimally complete* if it is complete and if the sum total of all the single gossips in all the multigossips in the sequence is no larger than the sum total of all the single gossips in all the multigossips in any other complete multigossip sequence. It is clear that if Σ is minimally complete, then the subgraph $\mathbb{N}_\Sigma$ it induces must be a minimal spanning tree. On the other hand, if Σ is nonredundant and $\mathbb{N}_\Sigma$ is a minimal spanning tree of $\mathbb{N}$, then Σ must be minimally complete.

C. Convergence

Roughly speaking, if over a period T a complete multigossip sequence has occurred, then each agent in the group will have been "in touch" with each other agent at least indirectly. It is not surprising then that complete multigossip subsequences over successive periods in an infinitely long sequence should be sufficient for all gossip variables in a gossiping process to converge to a common value. Prompted by this, let us call an infinite sequence of multigossips *repetitively complete* with period T if each successive subsequence of multigossips of length T in the sequence is complete. The following theorem implies that repetitive complete multigossip sequences converge exponentially fast.

Theorem 1: Let $M(1), M(2), M(3), \ldots$ denote the gossiping matrices corresponding to an infinite sequence of multigossips which is repetitively complete with period T. Suppose that the vector of gossip variables $x(t)$ evolves according to $x(t+1) = M(t)x(t),\ t \geq 1$. There exists a real nonnegative number $\lambda < 1$ such that for each initial value of $x(1)$, all n gossip variables converge to the average value

$$\frac{1}{n}\sum_{i=1}^{n} x_i(1)$$

as fast as λ^t converges to zero.

There are several different ways to prove this theorem; see, for example, [3], [6], [7], and [23]. It turns out that this theorem is a simple consequence of more technical results which are of interest in their own right and which will be stated and proved in Section IV.

The theorem also applies to more general gossiping algorithms in which simple averaging between agent pairs is replaced with averaging based on convex combinations. All that is required is that the averaging rule determines doubly stochastic matrices $M(\cdot)$ which take values in a *compact* set.

D. Tree Graphs

In graph theory, tree graphs (i.e., graphs without cycles) often lead to significant simplifications. This is also the case with gossiping. Let us note that a tree graph has the property that removal of any one edge (i, j) results in a disconnected graph.[2] This means that if $\mathbb{N}$ is a tree, a necessary condition for a finite sequence of gossips to be complete is that over the period during which the gossips occur, each agent must gossip with each of its neighbors at least once. It is clear that the converse is also true; i.e., if each agent gossips with each of its neighbors at least once during a given period, then the sequence of gossips which took place over that period must be complete. It is easy to see that in the case when $\mathbb{N}$ is a tree, a gossip sequence is minimally complete if and only if a gossip between each agent and each of its neighbors occurs exactly once in the sequence. Equivalently a gossip matrix G for a graph $\mathbb{N}$, which is a tree, is minimally complete if and only if G is a product of all of the single-gossip matrices associated with $\mathbb{N}$.

III. PERIODIC GOSSIPING

About the easiest way to guarantee a repetitively complete multigossip sequence is to use a protocol which generates an infinite multigossip sequence which on the one hand is "periodic" and on the other is complete on each successive period. Prompted by this, let us agree to call an infinite sequence of multigossips *periodic* with period T if each multigossip in the sequence occurs once every T time units; such a sequence is *periodically complete* if each subsequence consisting of T consecutive multigossips is complete. It is clear that any periodically complete multigossip sequence is repetitively complete. The converse of course is not true.

A. Convergence Rate

Corresponding to any T-periodic sequence of multigossips is an infinite sequence of gossip matrices; such a matrix sequence is *periodic* with period T in that each matrix within the sequence repeats itself every T time units. Suppose that $M(1), M(2), \ldots$ is such a T-periodic sequence. If $x(t+1) = M(t)x(t),\ t \geq 1$, it is clear that $x((i+1)T+1) = Nx(iT+1),\ i \geq 0$, where $N = M(T)M(T-1)\cdots M(1)$. Thus, $x(iT+1) = N^i x(1),\ i \geq 0$, which means that both the convergence and convergence rate of the gossip variables in a periodic multigossip sequence are completely determined by properties of N. Note that N is a doubly stochastic matrix because each of the matrices in the product defining it is doubly stochastic, and because the class of $n \times n$ doubly stochastic matrices is closed under multiplication. Now because N is stochastic, it has an eigenvalue at 1 and its spectral radius is 1 [27]. We are interested in the case when $\lim_{i\to\infty} N^i = (1/n)\mathbf{1}\mathbf{1}'$ which is clearly just when all eigenvalues other than the one eigenvalue with

[2]If this were not so then there would have to be at least two distinct paths between i and j which contradicts the requirement that a tree be acyclic.

value 1 have magnitudes strictly less than one. This is precisely the property of a complete gossip matrix.

Theorem 2: A gossip matrix is complete if and only if the magnitudes of all of its eigenvalues, with the exception of a single eigenvalue of value 1, are strictly less than 1.

A proof of this theorem will be given in Section IV-D.

For any doubly stochastic matrix S, let $\rho(S)$ denote the magnitude of the second largest eigenvalue (in magnitude) of S. It is clear that the rate at which x converges to $1y_{\text{avg}}$ is $\rho^{1/T}(N)$. In the case when $\mathbb{N}$ is a tree, $\rho(N)$ turns out to be the same for all minimally complete gossip matrices determined by $\mathbb{N}$. This somewhat surprising fact is a direct consequence of the following theorem which is the main result of [25].

Theorem 3: Let $\mathcal{E} = e_1, e_2, \ldots, e_k$ be the sequence of edges of $\mathbb{N}$ labeling a nonredundant, complete gossip sequence. Let $\mathbb{N}_{\mathcal{E}}$ denote the spanning subgraph of $\mathbb{N}$ whose edges are the edges in $\mathcal{E}$. Let $\mathcal{G}(\mathcal{E})$ be the group (under composition) consisting of the identity map on $\{1, 2, \ldots, k\}$ together with all maps $\pi : \{1, 2, \ldots, k\} \to \{1, 2, \ldots, k\}$ generated by all permutations which satisfy one of the following conditions:

1) π is a cyclic permutation of $\{1, 2, \ldots, k\}$;
2) π is that permutation of $\{1, 2, \ldots, k\}$ which for some $i < k$, interchanges i and $i + 1$ provided that in $\mathbb{N}_{\mathcal{E}}$, either
   a) e_i and e_{i+1} are not incident on the same vertex or
   b) e_i and e_{i+1} are incident on the same vertex but neither edge is contained in any cycle of $\mathbb{N}_{\mathcal{E}}$.

For each $\pi \in \mathcal{G}(\mathcal{E})$, let G_π denote the gossip matrix induced by the edge sequence $e_{\pi(1)}, e_{\pi(2)}, \ldots, e_{\pi(k)}$. Then, the characteristic polynomial of G_π is the same for all $\pi \in \mathcal{G}(\mathcal{E})$.

A proof of this theorem can be found in [25]. See also [21] for an alternative proof.

Note that if $\mathbb{N}_{\mathcal{E}}$ is a tree, then condition 2 in Theorem 3 implies that for any two successive integers i and $i + 1$ in $\{1, 2, \ldots, k\}$, there is a permutation in $\mathcal{G}(\mathcal{E})$ which interchanges i and $i + 1$. A simple induction thus proves that if $\mathbb{N}_{\mathcal{E}}$ is a tree, $\mathcal{G}(\mathcal{E})$ is the set of *all* permutations on $\{1, 2, \ldots, k\}$. Therefore, in this case, the characteristic polynomial of G_π is the same for all permutations of $\{1, 2, \ldots, k\}$.

Suppose that $\mathbb{N}$ is a tree. Then, because of completeness, $\mathbb{N}_{\mathcal{E}} = \mathbb{N}$, and therefore, $\mathcal{G}(\mathcal{E})$ is the set of all permutations on $\{1, 2, \ldots, k\}$. Thus, if $\mathbb{N}$ is a tree, Theorem 3 implies that $\rho(N)$ is the same for all minimal complete gossip matrices determined by $\mathbb{N}$. This conclusion is not implied by Theorem 3 if $\mathbb{N}$ is not a tree.

As stated, the theorem is only for sequences of single gossips. However, the same theorem also applies, with virtually the same proof, to sequences of multigossips. This is because for purposes of analysis, any multigossip can be viewed as a sequence of noninteracting single gossips arranged in any order.

B. Multigossiping

It is clear from the preceding that the rate at which the gossip variables of a periodically complete gossiping sequence converge depends not only on $\rho(N)$ but also on T. For example, suppose that $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \ldots, \gamma_T, \gamma_1, \gamma_2, \ldots$ is an infinite periodically complete gossip sequence with period T. Suppose in addition that $\gamma_1, \gamma_2, \gamma_3$ are noninteracting gossips. Then, these three gossips might be executed simultaneously as a multigossip $\{\gamma_1, \gamma_2, \gamma_3\}$, rather than sequentially, at the beginning of each period without in any way affecting the complete gossip matrix N corresponding to the original subsequence $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \ldots, \gamma_T$. In other words, rather than executing the T-periodic sequence $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \ldots, \gamma_T, \gamma_1, \gamma_2, \ldots$, the group could execute the periodic sequence $\{\gamma_1, \gamma_2, \gamma_3\}, \gamma_4, \ldots, \gamma_T, \{\gamma_1, \gamma_2, \gamma_3\}, \gamma_4, \ldots$ without changing the value of $\rho(N)$. The key point here is that this sequence has period $T - 2$ rather than T. Thus, by using multigossiping, the worst case convergence rate for this gossiping process would be reduced from $\rho^{1/T}(N)$ to $\rho^{1/(T-2)}(N)$. It is obvious that, in general, to get faster convergence, one would want to implement multigossiping sequences using the smallest number of distinct multigossips possible. For the case when $\mathbb{N}$ is a tree and the original subsequence $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \ldots, \gamma_T$ is minimally complete, we know that the order of the gossips in the sequence can be changed without changing $\rho(N)$. In this case, the minimal number of multigossips needed to implement the original sequence would be the same as the minimal number of colors needed to color the edges of $\mathbb{N}$ subject to the constraint that no two edges incident on any vertex have the same color, for edges of the same color would then correspond to those gossips which could be implemented together as a single multigossip. Edge coloring is a basic problem in graph theory [28]. The minimal number of colors required to color a graph subject to this constraint is called the *chromatic index*. Vizing's theorem states that the chromatic index of a neighbor graph $\mathbb{N}$ is either $\mathbf{d}$ or $\mathbf{d} + 1$ where $\mathbf{d}$ is the maximum vertex degree of $\mathbb{N}$ [29]. Moreover, if $\mathbb{N}$ is a tree, the chromatic index is $\mathbf{d}$ because of König's theorem [28]. In other words, if $\mathbb{N}$ is a tree with maximum vertex degree $\mathbf{d}$, it is possible to construct a periodic sequence of multigossips with period $T = \mathbf{d}$ which converges as fast as the sequence $\rho^{1/d}(N), \rho^{2/d}(N), \rho^{3/d}(N), \ldots$ converges to zero where N is any minimally complete gossip matrix for the graph.

IV. REQUEST-BASED GOSSIPING

Request-based gossiping is a gossiping process in which a gossip occurs between two agents whenever one of the two accepts a request to gossip placed by the other. The aim of this section is to discuss this process.

In a request-based gossiping process, a given agent i may gossip with one of its neighbor's at time t only if t is either an "event time" of agent i or an "event time" of its

neighbor which has made a request to gossip with agent i. By an *event time* of agent i is meant a time at which agent i may place a request to gossip with one of its neighbors. By an *event time interval* of agent i is meant the interval of time between two successive event times of agent i. For obvious reasons, we assume that the lengths of agent i's event time intervals are all bounded above by a finite positive number T_i. We write $\mathcal{T}_i$ for the set of event times of agent i and $\mathcal{T}$ for the union of the event time sequences of all n agents.

Conflicts leading to deadlocks can arise if an agent who has placed a request to gossip, at the same time receives a request to gossip from another agent. It is challenging to devise rules which resolve such conflicts while at the same time ensuring exponential convergence of the gossiping process. One way to avoid such conflicts is to assign event times offline so that no agent can receive a request to gossip at any of its own event times. There are several ways to do this which will be discussed below.

From time to time, agent i may have more than one neighbor to which it might be able to make a request to gossip with. Also from time to time, agent i may receive more than one request to gossip. While in such situations decisions about who to place requests with or whose request to accept can be randomized, in this paper, we will examine only completely deterministic strategies. To do this we will assume that each agent has ordered its neighbors in $\mathcal{N}_i$ according to some priorities so when a choice occurs between neighbors, agent i will always choose the one with highest priority.

Consider first the situation when the event times of each agent and each agent's neighbor priorities are chosen offline and are fixed throughout the gossiping process. Assume that the event times are chosen so that no agent can receive a request to gossip at any of its own event times. Our aim is to show that this arrangement can be problematic. The following protocol illustrates this.

Protocol I: At each event time $t \in \mathcal{T}$ the following rules apply for each $i \in \{1, 2, \ldots, n\}$.

1) If $t \in \mathcal{T}_i$, agent i places a request to gossip with that neighbor whose priority is the highest.
2) If $t \notin \mathcal{T}_i$, agent i does not place a request to gossip.
3) Each agent i receiving one or more requests to gossip must gossip with that requesting neighbor whose priority is the highest.
4) If $t \notin \mathcal{T}_i$ and agent i does not receive a request to gossip, it does not gossip.

The following example shows that this simple strategy will not necessarily lead to a consensus. Suppose that $\mathbb{N}$ is a path graph with edges $(a, b), (b, c), (c, d)$. Assume that agents a and b have distinct event times and that agents a and c have the same event times as do agents b and d; note that this guarantees that no agent can receive a request to gossip at any of its own event times. To avoid ambiguities in decision making, suppose that agent b assigns a higher priority to a than to c and agent c assigns a higher priority to d than to b. Let t be an event time of agents a and c. Then, at this time a places a request to gossip with b and c places a request to gossip with d. Since b and d receive no other requests, gossips take place between a and b and between c and d. Alternatively, if t is an event time of agents b and d, then at this time, b places a request to gossip with a and d places a request to gossip with c. Since a and c receive no other requests, gossips again take place between a and b and between c and d. Thus, under no conditions is there ever a gossip between b and c, so the gossiping process will never reach a consensus. The reader may wish to verify that simply changing the priorities will not rectify this situation: For any choice of priorities, there will always be at least one gossip needed to reach a consensus, which will not take place.

The preceding example illustrates that fixed priorities can present problems. In what follows we take an alternative approach.

In the light of Theorem 1 it is of interest to consider gossiping protocols which generate repetitively complete gossip sequences. Towards this end, let us agree to say that an agent i has completed a *round* of gossiping after it has gossiped with each neighbor in $\mathcal{N}_i$ at least once. Thus, a finite gossiping sequence for the entire group which has occurred over an interval of length T will be complete if over the same period each agent in the group completes a round. In fact in the case when $\mathbb{N}$ is a tree, the only way such a sequence could be complete is if over the same period each agent in the group completes a round.

For the protocols which follow it will be necessary for each agent i to keep track of where it is in a particular round. To do this, agent i makes use of a recursively updated *neighbor queue* $\mathbf{q}_i(t)$ where $\mathbf{q}_i(\cdot)$ is a function from $\mathcal{T}$ to the set of all possible lists of the m_i labels in $\mathcal{N}_i$, the neighbor set of agent i. Roughly speaking, $\mathbf{q}_i(t)$ is a list of the labels of the neighbors of agent i at time t which defines the queue of neighbors at time t which are in line to gossip with agent i. The updating of $\mathbf{q}_i(t)$ is straightforward: If neighbor j gossips with agent i at time t, the updated queue $\mathbf{q}_i(t+1)$ is obtained by moving agent j's label from its current position in $\mathbf{q}_i(t)$, to the end of the queue. If on the other hand, agent i does not gossip at time t, $\mathbf{q}_i(t+1) = \mathbf{q}_i(t)$.

A. Protocols

As noted earlier, it is helpful to have event time assignments which guarantee that no agent can receive a request to gossip at any of its own event times. One easy way to accomplish this is to use event time assignments which satisfy the following assumption.

Distinct Event Times Assumption: For each distinct pair of integers i and j in $\{1, 2, \ldots, n\}$, $\mathcal{T}_i$ and $\mathcal{T}_j$ are disjoint sets.

Note that the assumption implies that at any fixed event time in $\mathcal{T}$, only one agent can receive a request to

gossip, and moreover, that this agent will receive exactly one such request. A simple protocol which ensures exponential convergence under this assumption is as follows.

Protocol II: Suppose that the distinct event times assumption holds. At each event time $t \in \mathcal{T}$, the following rules apply for each $i \in \{1, 2, \ldots, n\}$.

1) If $t \in \mathcal{T}_i$, agent i places a request to gossip with that neighbor whose label is in the front of the queue $\mathbf{q}_i(t)$.
2) If $t \notin \mathcal{T}_i$, agent i does not place a request to gossip.
3) Each agent i receiving a request to gossip must gossip with the neighbor placing the request.
4) If $t \notin \mathcal{T}_i$ and agent i does not receive a request to gossip, it does not gossip.

The behavior of Protocol II can be easily explained as follows. For $i \in \{1, 2, \ldots, n\}$, let d_i be the number of neighbors of agent i, or equivalently, the degree of vertex i in $\mathbb{N}$. It is clear that with the distinct event times protocol, each agent i will complete a round in a time interval containing no more than d_i event time intervals of agent i. The length of the time interval large enough to contain d_i successive event time intervals of agent i for all $i \in \{1, 2, \ldots, n\}$ is the maximum of the times $T_i d_i$, $i \in \{1, 2, \ldots, n\}$. Thus, the sequence of gossips which occur on an interval of this length must necessarily be complete. We have proved the following.

Proposition 1: Suppose that the distinct event times assumption holds and that all agents in the group adhere to Protocol II. Then, the infinite sequence of gossips generated will be repetitively complete with period

$$T = \max_i (d_i T_i).$$

One shortcoming of Protocol II is that it does not allow for multigossiping. Another is that the distinct event times assumption on which the protocol depends is somewhat stringent. It is possible to relax this assumption and still devise a protocol which ensures exponential convergence. The relaxed assumption is as follows.

Distinct Neighbor Event Times Assumption: For each $i \in \{1, 2, \ldots, n\}$ and each $j \in \mathcal{N}_i$, $\mathcal{T}_i$ and $\mathcal{T}_j$ are disjoint sets.

Thus, if this assumption holds, the event times of each agent are distinct from the event times of all of its neighbors. The assignment of event times which satisfy this assumption is mathematically identical to the classic "vertex coloring problem" from graph theory [28]. Note that the distinct neighbor event times assumption stipulates that no two adjacent vertices on the neighbor graph $\mathbb{N}$ can have the same event times. The rule defining vertex coloring of $\mathbb{N}$ stipulates that no two adjacent vertices can have the same color. The least number of different colors required to vertex color $\mathbb{N}$ is called the *chromatic number* of $\mathbb{N}$ [28]. Brooks' theorem states that this number is bounded above by the maximum degree of $\mathbb{N}$, except for complete graphs and for graphs with cycles of odd length in which cases the bound is one plus the maximum degree of $\mathbb{N}$ [30]. Thus, in all cases the largest number of distinct event time sequences which would need to be assigned to $\mathbb{N}$ to satisfy the distinct neighbor event times assumption is no greater than one plus the maximum vertex degree of $\mathbb{N}$.

Under the distinct neighbor event times assumption, it is possible to ensure exponential convergence with the following protocol which is a refinement of Protocol II.

Protocol III: Suppose that the distinct neighbor event times assumption holds. At each event time $t \in \mathcal{T}$ the following rules apply for each $i \in \{1, 2, \ldots, n\}$.

1) If $t \in \mathcal{T}_i$, agent i places a request to gossip with that neighbor whose label is at the front of the queue $\mathbf{q}_i(t)$.
2) If $t \notin \mathcal{T}_i$, agent i does not place a request to gossip.
3) Each agent i receiving one or more requests to gossip must gossip with that requesting neighbor whose label is closest to the front of the queue $\mathbf{q}_i(t)$.
4) If $t \notin \mathcal{T}_i$ and agent i does not receive a request to gossip, it does not gossip.

Just as with the distinct event times protocol, it is possible to derive a worst case bound on the time it takes for all n agents to complete a round of gossiping. As a first step towards this end, fix i and suppose that at some given event time $t_0 \in \mathcal{T}_i$, j is the leading label in the queue $\mathbf{q}_i(t_0)$. According to the preceding protocol, from this event time forward agent i must repeatedly place requests with agent j to gossip at successive event times in $\mathcal{T}_i$, until gossiping between the two takes place. Meanwhile, at these same event times, neighbor j will be receiving requests to gossip from neighbor i and possible some other neighbors. In the worst case, when label i is at the end of $\mathbf{q}_j(t_0)$ at time t_0, it will take at most d_j event time intervals of agent i for label i to advance to the front of agent j's queue. This means that agents i and j are guaranteed to gossip at least once within any time interval containing no more than d_j event time intervals of agent i. If agent i has only one neighbor, then the round is complete in at most $T_i d_j$ time units. On the other hand, if agent i has more than one neighbor, agent i then begins to place requests to gossip with the agent whose label k was second from the front in the queue $\mathbf{q}_i(t_0)$. But at this time label i might be, in the worst case, at the end of the queue for agent k. By the same reasoning as before, it will take at worst an additional d_k successive event time intervals of agent i for agents i and k to gossip. In other words, agent i is guaranteed to have gossiped at least once with both agent j and agent k within any time interval containing no more than

$d_j + d_k$ event time intervals of agent i. By repeating this argument for all labels in the queue $\mathbf{q}_i(t_0)$, one reaches the conclusion that agent i is guaranteed to complete a round of gossiping with all of its neighbors in any time interval containing $\sum_{j\in\mathcal{N}_i} d_j$ event time intervals of agent i. An upper bound on the length of any such interval is thus $T_i \sum_{j\in\mathcal{N}_i} d_j$.

From the preceding it is clear that within any interval of time containing either $\sum_{j\in\mathcal{N}_i} d_j$ event times of agent i or $\sum_{j\in\mathcal{N}_k} d_j$ event time intervals of agent k, neighbors i and k will gossip at least once; thus neighbors i and k will gossip at least once in any time interval of length $\min\{T_i \sum_{j\in\mathcal{N}_i} d_j, T_k \sum_{j\in\mathcal{N}_k} d_j\}$. The maximum of this amount of time over all agent pairs is thus an upper bound on the amount of time it takes for all neighbor pairs to gossip at least once. But completeness of a gossip sequence is assured if the sequence contains a gossip for each possible neighbor pair. We have therefore proved the following proposition.

Proposition 2: Let $\mathcal{E}$ denote the set of all edges (i, j) in $\mathbb{N}$. Suppose that the distinct neighbor event times assumption holds and that all agents in the group adhere to Protocol III. Then, the infinite sequence of gossips generated will be repetitively complete with period

$$T = \max_{(i,k)\in\mathcal{E}} \min\left\{ T_i \sum_{j\in\mathcal{N}_i} d_j,\ T_k \sum_{j\in\mathcal{N}_k} d_j \right\}.$$

A disadvantage of Protocol III is that it requires the distinct neighbor event times assumption. This assumption can only be satisfied by offline assignment of event times for each agent, and in some applications such an offline assignment may be undesirable. In a recent doctoral thesis [13], a clever gossiping protocol is proposed which does not require the distinct neighbor event times assumption. The protocol avoids deadlocks and achieves consensus exponentially fast. A disadvantage of this protocol is that it requires each agent to obtain the values of *all* of its neighbors' gossip variables at each clock time. Thus, if communication cost is an important issue, this protocol may not be satisfactory even though only local information is required. By exploiting one of the key ideas in [13] together with the notion of an agent's neighbor queue $\mathbf{q}_i(t)$ defined earlier, it is possible to obtain a gossiping protocol which also avoids deadlocks and achieves consensus exponentially fast but without requiring each agent to obtain the value of more than one of its neighbors' gossip variables at each clock time. Our aim below is to outline this protocol.

Protocol IV: In the following, agent i's *preferred neighbor* at time t is that agent whose label $i^*(t)$ is in the front of the queue $\mathbf{q}_i(t)$. Between clock times t and $t+1$ each agent i performs the steps enumerated below in the order indicated. Although the agents' actions need not be precisely synchronized, it is understood that for each $k \in \{1, 2, 3\}$ all agents complete step k before any embark on step $k+1$.

1) First Transmission: Agent i places a request to gossip with its preferred neighbor by sending both its label i and its gossip value $x_i(t)$ to agent $i^*(t)$. At the same time agent i receives requests to gossip (i.e., the labels and corresponding gossip values) from all of those neighbors which have agent i as their current preferred neighbor. Let $\mathcal{R}_i(t)$ denote the set of labels of these requesting neighbors.
2) Second Transmission: Agent i sends its current gossip value $x_i(t)$ to those neighbors which have agent i as their current preferred neighbor, namely the neighbors of agent i with labels in $\mathcal{R}_i(t)$.
3) Acceptances:
   a) If agent i has not placed a request to gossip but has received at least one request to gossip, then agent i sends an acceptance to that particular requesting neighbor whose label is closest to the front of the queue $\mathbf{q}_i(t)$.
   b) If agent i has either placed a request to gossip or has not received any requests to gossip, then agent i does not send out an acceptance.
4) Gossip variable and queue updates:
   a) If agent i either sends an acceptance to or receives an acceptance from neighbor j, then agent i gossips with neighbor j by setting

$$x_i(t+1) = \frac{x_i(t) + x_j(t)}{2}.$$

   Agent i updates its queue by moving j, and any labels $k \in \{i^*(t)\} \cup \mathcal{R}_i(t)$ for which $x_k(t) = x_i(t)$ from their current positions in $\mathbf{q}_i(t)$ to the end of the queue while maintaining their relative order.
   b) If agent i has not sent out an acceptance or received one, then agent i does not update the value of $x_i(t)$. In addition, $\mathbf{q}_i(t)$ is updated by moving any labels $k \in \{i^*(t)\} \cup \mathcal{R}_i(t)$ for which $x_k(t) = x_i(t)$ from their current positions in $\mathbf{q}_i(t)$ to the end of the queue while maintaining their relative order.

It is possible to show that every gossip sequence generated by this protocol is repetitively complete with period no greater than the number of edges of $\mathbb{N}$ [31]. It follows from Theorem 1 that any sequence of gossip vectors generated by this protocol is exponentially convergent.

1) Convergence Rate: Recall that a gossiping sequence is repetitively complete with period T if each successive subsequence of gossips of length T in the sequence is complete; and if, in addition, each gossip in the sequence repeats once every T time units, the sequence is periodic with period T. As was noted in Section III-A, for a repetitively complete sequence of gossiping matrices $M(1)$, $M(2), \ldots$ which is T-periodic, the convergence rate of the product $M(T)M(T-1)\cdots M(1)$ as $t \to \infty$ is determined by T and by the eigenvalue of $N = M(T)M(T-1)\cdots M(1)$ which is second largest in magnitude. For gossiping sequences which are repetitively complete but not periodic this is no longer true. Such sequences are closely related to what are called "nonhomogeneous Markov chains" for which there is a substantial literature [26]. Notwithstanding this, the following question remains. What determines the convergence rate of a repetitively complete gossip sequence which is not necessarily periodic? This is the question which will be considered next. We will tackle the question in two steps. First, in Section IV-B, we will discuss certain relevant basic properties of stochastic matrices. Then, in Section IV-C, we will study several types of "seminorms" appropriate to the analysis of nonhomogeneous Markov chains. Finally, in Section IV-D, we will show that a certain seminorm provides exactly what is needed to characterize the convergence rate of a repetitively complete gossip sequence.

B. Stochastic Matrices

Since gossip matrices are stochastic matrices, a natural starting point for the study of convergence rates of gossiping sequences is a review of some of the basic concepts associated with stochastic matrices. We begin with the idea of a graph of a stochastic matrix.

1) Graph of a Stochastic Matrix: Many properties of a stochastic matrix can be usefully described in terms of an associated directed graph determined by the matrix. The *graph* of nonnegative matrix $M \in \mathbb{R}^{n\times n}$, written $\gamma(M)$, is a directed graph on n vertices with an arc from vertex i to vertex j just in case $m_{ji} \neq 0$; if (i,j) is such an arc, we say that i is a *neighbor* of j and that j is an *observer* of i. Thus, $\gamma(M)$ is that directed graph whose adjacency matrix is the transpose of the matrix obtained by replacing all nonzero entries in M with ones.

2) Connectivity: There are various notions of connectivity which are useful in the study of the convergence of products of stochastic matrices. Perhaps the most familiar of these is the idea of "strong connectivity." A directed graph is *strongly connected* if there is a directed path between each pair of distinct vertices. A directed graph is *weakly connected* if there is an undirected path between each pair of distinct vertices. There are other notions of connectivity which are also useful in this context. To define several of them, let us agree to call a vertex i of a directed graph $\mathbb{G}$, a *root* of $\mathbb{G}$ if for each other vertex j of $\mathbb{G}$, there is a directed path from i to j. Thus, i is a root of $\mathbb{G}$, if it is the root of a directed spanning tree of $\mathbb{G}$. We will say that $\mathbb{G}$ is *rooted at* i if i is in fact a root. Thus, $\mathbb{G}$ is rooted at i just in case each other vertex of $\mathbb{G}$ is *reachable* from vertex i along a directed path within the graph. $\mathbb{G}$ is *strongly rooted at* i if each other vertex of $\mathbb{G}$ is reachable from vertex i along a directed path of length 1. Thus, $\mathbb{G}$ is strongly rooted at i if i is a neighbor of every other vertex in the graph. By a *rooted graph* $\mathbb{G}$ is meant a directed graph which possesses at least one root. A *strongly rooted graph* is a graph which has at least one vertex at which it is strongly rooted. Note that a nonnegative matrix $M \in \mathbb{R}^{n\times n}$ has a strongly rooted graph if and only if it has a positive column. Note that every strongly connected graph is rooted and every rooted graph is weakly connected. The converse statements are false. In particular there are weakly connected graphs which are not rooted and rooted graphs which are not strongly connected.

3) Composition: Since we will be interested in products of stochastic matrices, we will be interested in graphs of such products and how they are related to the graphs of the matrices comprising the products. For this we need the idea of "composition" of graphs. Let $\mathbb{G}_p$ and $\mathbb{G}_q$ be two directed graphs with vertex set $\mathcal{V}$. By the *composition* of $\mathbb{G}_p$ with $\mathbb{G}_q$, written $\mathbb{G}_q \circ \mathbb{G}_p$, is meant the directed graph with vertex set $\mathcal{V}$ and arc set defined in such a way so that (i,j) is an arc of the composition just in case there is a vertex k such that (i,k) is an arc of $\mathbb{G}_p$ and (k,j) is an arc of $\mathbb{G}_q$. Thus, (i,j) is an arc in $\mathbb{G}_q \circ \mathbb{G}_p$ if and only if i has an observer in $\mathbb{G}_p$ which is also a neighbor of j in $\mathbb{G}_q$. Note that composition is an associative binary operation; because of this, the definition extends unambiguously to any finite sequence of directed graphs $\mathbb{G}_1, \mathbb{G}_2, \ldots, \mathbb{G}_k$ with the same vertex set.

Composition and matrix multiplication are closely related. In particular, the graph of the product of two nonnegative matrices $M_1, M_2 \in \mathbb{R}^{n\times n}$ is equal to the composition of the graphs of the two matrices comprising the product. In other words, $\gamma(M_2M_1) = \gamma(M_2) \circ \gamma(M_1)$.

If we focus exclusively on graphs with self-arcs at all vertices, more can be said. In this case, the definition of composition implies that the arcs of both $\mathbb{G}_p$ and $\mathbb{G}_q$ are arcs of $\mathbb{G}_q \circ \mathbb{G}_p$; the converse is false. The definition of composition also implies that if $\mathbb{G}_p$ has a directed path from i to k and $\mathbb{G}_q$ has a directed path from k to j, then $\mathbb{G}_q \circ \mathbb{G}_p$ has a directed path from i to j. These implications are consequences of the requirement that the vertices of the graphs in question all have self-arcs. It is worth emphasizing that the union of the arc sets of a sequence of graphs $\mathbb{G}_1, \mathbb{G}_2, \ldots, \mathbb{G}_k$ with self-arcs must be contained in the arc set of their composition. However, the converse is not true in general and it is for this reason that composition rather than union proves to be the more useful concept for our purposes.

4) Convergability: It is of obvious interest to have a clear understanding of what kinds of stochastic matrices within an infinite product guarantee that the infinite product converges. There are many ways to address this issue and many existing results. Here we focus on just one issue.

Let $\mathcal{S}$ denote the set of all stochastic matrices in $\mathbb{R}^{n\times n}$ with positive diagonal entries. Call a compact subset $\mathcal{M} \subset \mathcal{S}$ *convergable* if for each infinite sequence of matrices $M_1, M_2, M_3, \ldots$ from $\mathcal{M}$, the sequence of products $M_1, M_2M_1, M_3M_2M_1, \ldots$ converges exponentially fast to a matrix of the form $\mathbf{1}c$. Convergability can be characterized as follows.

Theorem 4: Let $\mathcal{R}$ denote the set of all matrices in $\mathcal{S}$ with rooted graphs. Then, a compact subset $\mathcal{M} \subset \mathcal{S}$ is convergable if and only if $\mathcal{M} \subset \mathcal{R}$.

The theorem implies that $\mathcal{R}$ is the largest subset of $n \times n$ stochastic matrices with positive diagonal entries whose compact subsets are all convergable. $\mathcal{R}$ itself is not convergable because it is not closed and thus not compact.

Proof of Theorem 4: The fact that any compact subset of $\mathcal{R}$ is convergable is more or less well known from the work reported in [32]; the statement also follows from Proposition 11 of [33]. To prove the converse, suppose that $\mathcal{M} \subset \mathcal{S}$ is convergable. Then, by continuity, every sufficiently long product of matrices from $\mathcal{M}$ must be a matrix with a positive column. Therefore, the graph of every sufficiently long product of matrices from $\mathcal{M}$ must be strongly rooted. It follows from Proposition 5 of [33] that $\mathcal{M}$ must be a subset of $\mathcal{R}$. ■

Although doubly stochastic matrices are stochastic, convergability for classes of doubly stochastic matrices has a different characterization than it does for classes of stochastic matrices. Let $\mathcal{D}$ denote the set of all doubly stochastic matrices in $\mathcal{S}$. In the following, we will prove Theorem 5.

Theorem 5: Let $\mathcal{W}$ denote the set of all matrices in $\mathcal{D}$ with weakly connected graphs. Then, a compact subset $\mathcal{M} \subset \mathcal{D}$ is convergable if and only if $\mathcal{M} \subset \mathcal{W}$.

The theorem implies that $\mathcal{W}$ is the largest subset of $n \times n$ doubly stochastic matrices with positive diagonal entries whose compact subsets are all convergable. Like $\mathcal{R}$, $\mathcal{W}$ is not convergable because it is not compact.

An interesting set of stochastic matrices in $\mathcal{S}$ whose compact subsets are known to be convergable is the set of all "scrambling matrices." A matrix $S \in \mathcal{S}$ is *scrambling* if for each distinct pair of integers i and j, there is a column k of S for which s_{ik} and s_{jk} are both nonzero [26]. In graph theoretic terms S is a scrambling matrix just in case its graph is "neighbor shared" where by *neighbor shared* we mean that each distinct pair of vertices in the graph share a common neighbor [33]. Convergability of compact subsets of scrambling matrices is tied up with the concept of the coefficient of ergodicity [26] which for a given stochastic matrix $S \in \mathcal{S}$ is defined by

$$\tau(S) = \frac{1}{2}\max_{i,j}\sum_{k=1}^{n}|s_{ik} - s_{jk}|. \tag{1}$$

It is known that $0 \le \tau(S) \le 1$ for all $S \in \mathcal{S}$ and that

$$\tau(S) < 1 \tag{2}$$

if and only if S is a scrambling matrix. It is also known that

$$\tau(S_2S_1) \le \tau(S_2)\tau(S_1), \qquad S_1, S_2 \in \mathcal{S}. \tag{3}$$

It can be shown that (2) and (3) are sufficient conditions to ensure that any compact subset of scrambling matrices is convergable. But $\tau(\cdot)$ has another role. It provides a worst case convergence rate for any infinite product of scrambling matrices from a given compact set $\mathcal{C} \subset \mathcal{S}$. In particular, it can be easily shown that as $i \to \infty$, any product $S_iS_{i-1}\cdots S_2S_1$ of scrambling matrices $S_i \in \mathcal{C}$ converges to a matrix of the form $\mathbf{1}c$ as fast as λ^i converges to zero where

$$\lambda = \max_{S\in\mathcal{C}}\tau(S).$$

This preceding discussion suggests the following question. Can analogs of the coefficient of ergodicity satisfying formulas like (2) and (3) be found for the set of stochastic matrices with rooted graphs or perhaps for the set of doubly stochastic matrices with weakly connected graphs? In the following, we will provide a partial answer to this question for the case of stochastic matrices and a complete answer for the case of doubly stochastic matrices. Our approach will be to appeal to certain types of seminorms of stochastic matrices.

C. Seminorms

Let $\|\cdot\|_p$ be the induced p-norm on $\mathbb{R}^{m\times n}$. We will be interested in $p = 1, 2, \infty$. Note that for a nonnegative matrix A

$$\begin{aligned}\|A\|_1 &= \max \text{ column sum } A\\ \|A\|_2 &= \sqrt{\mu(A'A)}\\ \|A\|_\infty &= \max \text{ row sum } A\end{aligned}$$

where $\mu(A'A)$ is the largest eigenvalue of $A'A$; that is, the square of the largest singular value of A. For $M \in \mathbb{R}^{m\times n}$

define

$$|M|_p = \min_{c\in\mathbb{R}^{1\times n}} \|M - \mathbf{1}c\|_p.$$

As defined, $|\cdot|_p$ is nonnegative and $|M|_p \le \|M\|_p$; clearly $|\mu M|_p = |\mu||M|_p$ for all real numbers μ so $|\cdot|_p$ is "positively homogeneous" [27]. Let M_1 and M_2 be matrices in $\mathbb{R}^{m\times n}$ and let c_0, c_1, and c_2 denote values of c which minimize $\|M_1 + M_2 - \mathbf{1}c\|_p$, $\|M_1 - \mathbf{1}c\|_p$, and $\|M_2 - \mathbf{1}c\|_p$, respectively. Note that

$$\begin{aligned}|M_1 + M_2|_p &= \|M_1 + M_2 - \mathbf{1}c_0\|_p\\ &\le \|M_1 + M_2 - \mathbf{1}(c_1 + c_2)\|_p\\ &\le \|M_1 - \mathbf{1}c_1\|_p + \|M_2 - \mathbf{1}c_2\|_p\\ &= |M_1|_p + |M_2|_p.\end{aligned}$$

Thus, the triangle inequality holds. These properties mean that $|\cdot|_p$ is a *seminorm*. $|\cdot|_p$ behaves much like a norm. For example, if N is a submatrix of M, then $|N|_p \le |M|_p$. However, $|\cdot|_p$ is not a norm because $|M|_p = 0$ does not imply $M = 0$; rather it implies that $M = \mathbf{1}c$ for some row vector c which minimizes $\|M - \mathbf{1}c\|_p$. For our purposes, $|\cdot|_p$ has a particularly important property.

Lemma 1: Suppose $\mathcal{M}$ is a subset of $\mathbb{R}^{n\times n}$ such that $M\mathbf{1} = \mathbf{1}$ for all $M \in \mathcal{M}$. Then

$$|M_2M_1|_p \le |M_2|_p|M_1|_p. \tag{4}$$

We say that $|\cdot|_p$ is *submultiplicative* on $\mathcal{M}$.

Proof of Lemma 1: Let c_0, c_1, and c_2 denote values of c which minimize $\|M_2M_1 - \mathbf{1}c\|_p$, $\|M_1 - \mathbf{1}c\|_p$, and $\|M_2 - \mathbf{1}c\|_p$, respectively. Then

$$\begin{aligned}|M_2M_1|_p &= \|M_2M_1 - \mathbf{1}c_0\|_p\\ &\le \|M_2M_1 - \mathbf{1}(c_2M_1 + c_1 - c_2\mathbf{1}c_1)\|_p\\ &= \|M_2M_1 - \mathbf{1}c_2M_1 - M_2\mathbf{1}c_1 + \mathbf{1}c_2\mathbf{1}c_1\|_p\\ &= \|(M_2 - \mathbf{1}c_2)(M_1 - \mathbf{1}c_1)\|_p\\ &\le \|(M_2 - \mathbf{1}c_2)\|_p\|(M_1 - \mathbf{1}c_1)\|_p\\ &= |M_2|_p|M_1|_p.\end{aligned}$$

Thus, (4) is true. ■

We say that $M \in \mathbb{R}^{n\times n}$ is *semicontractive* in the p-norm if $|M|_p < 1$. In view of Lemma 1, the product of semicontractive matrices in $\mathcal{M}$ is thus semicontractive. The importance of these ideas lies in the following fact.

Proposition 3: Suppose $\mathcal{M}$ is a subset of $\mathbb{R}^{n\times n}$ such that $M\mathbf{1} = \mathbf{1}$ for all $M \in \mathcal{M}$. Let p be fixed and let $\bar{\mathcal{M}}$ be a compact set of semicontractive matrices in $\mathcal{M}$. Let

$$\lambda = \sup_{\bar{\mathcal{M}}} |M|_p.$$

Then, for each infinite sequence of matrices $M_i \in \bar{\mathcal{M}}$, $i \in \{1, 2, \ldots\}$, the matrix product

$$M_iM_{i-1}\cdots M_1$$

converges as $i \to \infty$ as fast as λ^i converges to zero, to a rank one matrix of the form $\mathbf{1}\bar{c}$.

Proof of Proposition 3: Clearly $|M|_p \le \lambda$, $M \in \bar{\mathcal{M}}$. Moreover, $\lambda < 1$ because each $M \in \bar{\mathcal{M}}$ is semicontractive and because $\bar{\mathcal{M}}$ is compact. Write $M_i = \mathbf{1}c_i + T_i$, $i \ge 1$, where c_i is a value of c which minimizes $\|M_i - \mathbf{1}c\|$. For $i \ge 1$ set $X_i = M_iM_{i-1}\cdots M_1$ and $Y_i = T_iT_{i-1}\cdots T_1$. Clearly $|M_i|_p = \|T_i\|_p$, $i \ge 1$, so

$$\|Y_i\|_p \le \lambda^i, \qquad i \ge 1. \tag{5}$$

A simple computation yields

$$X_k = Y_k + \sum_{i=1}^{k} \mathbf{1}c_iY_{i-1}, \qquad k \ge 1 \tag{6}$$

where $Y_0 = I$. Note also that because of (5), Y_k tends to zero as $k \to \infty$. We claim that the sequence $\sum_{i=1}^{k} \mathbf{1}c_iY_{i-1}$, $k \ge 1$, has a limit. To prove that this is so it is enough to show that $\sum_{i=1}^{k} \mathbf{1}c_iY_{i-1}$, $k \ge 1$, is a Cauchy sequence. Towards this end observe that

$$\sum_{i=1}^{j+k} \mathbf{1}c_iY_{i-1} - \sum_{i=1}^{k} \mathbf{1}c_iY_{i-1} = \sum_{i=k+1}^{j+k} \mathbf{1}c_iY_{i-1}, \qquad j \ge 1.$$

Moreover

$$\|\mathbf{1}c_iY_{i-1}\|_p \le d\lambda^{i-1}, \qquad i \ge 1$$

where $d = \sup_{i\geq 1} \|\mathbf{1}c_i\|_p$. Therefore, for $j \geq 1$

$$\left\| \sum_{i=k+1}^{j+k} \mathbf{1}c_i Y_{i-1} \right\|_p \leq d \sum_{i=k+1}^{j+k} \lambda^{i-1} = d\lambda^k \sum_{s=1}^{j} \lambda^{s-1} \leq d\lambda^k \sum_{s=1}^{\infty} \lambda^{s-1} = d\frac{\lambda^k}{1-\lambda}.$$

Therefore

$$\left\| \sum_{i=1}^{j+k} \mathbf{1}c_i Y_{i-1} - \sum_{i=1}^{k} \mathbf{1}c_i Y_{i-1} \right\|_p \leq d\frac{\lambda^k}{1-\lambda}, \qquad j,k \geq 1 \quad (7)$$

which shows that $\sum_{i=1}^{k} \mathbf{1}c_i Y_{i-1}$, $k \geq 1$, is a Cauchy sequence. Therefore, the sequence $\sum_{i=1}^{k} \mathbf{1}c_i Y_{i-1}$, $k \geq 1$, has a limit which we denote by $\mathbf{1}\bar{c}$.

Note next that for $j, k \geq 1$

$$\left\| \sum_{i=1}^{k} \mathbf{1}c_i Y_{i-1} - \mathbf{1}\bar{c} \right\|_p = \left\| \sum_{i=1}^{j+k} \mathbf{1}c_i Y_{i-1} - \mathbf{1}\bar{c} - \left(\sum_{i=1}^{j+k} \mathbf{1}c_i Y_{i-1} - \sum_{i=1}^{k} \mathbf{1}c_i Y_{i-1} \right) \right\|_p$$

so

$$\left\| \sum_{i=1}^{k} \mathbf{1}c_i Y_{i-1} - \mathbf{1}\bar{c} \right\|_p \leq \left\| \sum_{i=1}^{j+k} \mathbf{1}c_i Y_{i-1} - \mathbf{1}\bar{c} \right\|_p + \left\| \sum_{i=1}^{j+k} \mathbf{1}c_i Y_{i-1} - \sum_{i=1}^{k} \mathbf{1}c_i Y_{i-1} \right\|_p.$$

In view of (7)

$$\left\| \sum_{i=1}^{k} \mathbf{1}c_i Y_{i-1} - \mathbf{1}\bar{c} \right\|_p \leq \left\| \sum_{i=1}^{j+k} \mathbf{1}c_i Y_{i-1} - \mathbf{1}\bar{c} \right\|_p + d\frac{\lambda^k}{1-\lambda}.$$

But $\| \sum_{i=1}^{j+k} \mathbf{1}c_i Y_{i-1} - \mathbf{1}\bar{c} \|_p$ tends to zero as $j \to \infty$ so

$$\left\| \sum_{i=1}^{k} \mathbf{1}c_i Y_{i-1} - \mathbf{1}\bar{c} \right\|_p \leq d\frac{\lambda^k}{1-\lambda}, \qquad k \geq 1. \quad (8)$$

To proceed observe that (6) implies that

$$\|X_k - \mathbf{1}\bar{c}\|_p \leq \|Y_k\|_p + \left\| \sum_{i=1}^{k} \mathbf{1}c_i Y_{i-1} - \mathbf{1}\bar{c} \right\|_p.$$

From this, (5), and (8), it follows that

$$\|X_k - \mathbf{1}\bar{c}\|_p \leq \lambda^k \left(1 + \frac{d}{1-\lambda}\right), \qquad k \geq 1.$$

This completes the proof. ■

1) Case $p = 1$: We now consider in more detail the case when $p = 1$. For this case, it is possible to derive an explicit formula for the seminorm $|M|_1$ of a nonnegative matrix $M \in \mathbb{R}^{n\times n}$.

Proposition 4: Let q be the unique integer quotient of n divided by 2. Let $M \in \mathbb{R}^{n\times n}$ be a nonnegative matrix. Then

$$|M|_1 = \max_{j\in\{1,2,\ldots,n\}} \left\{ \sum_{i\in\mathcal{L}_j} m_{ij} - \sum_{i\in\mathcal{S}_j} m_{ij} \right\}$$

where $\mathcal{L}_j$ and $\mathcal{S}_j$ are, respectively, the row indices of the q largest and q smallest entries in the jth column of M.

This result is a direct consequence of the following lemma and the definition of $|\cdot|_1$.

Lemma 2: Let q denote the unique integer quotient of n divided by 2. Let y be a nonnegative n-vector and write $\mathcal{L}$ and $\mathcal{S}$ for the row indices of the q largest and q smallest entries in y, respectively. Then

$$|y|_1 = \sum_{i\in\mathcal{L}} y_i - \sum_{i\in\mathcal{S}} y_i$$

where y_i is the ith entry in y.

Proof of Lemma 2: Let a denote the n-vector whose entries $a_1, a_2, \ldots, a_n$ are the entries of y relabeled so that $a_1 \leq a_2 \leq \cdots \leq a_n$. Then

$$\sum_{i\in\mathcal{L}} y_i = \sum_{i>(q+r)} a_i \quad \text{and} \quad \sum_{i\in\mathcal{S}} y_i = \sum_{i\leq q} a_i$$

where r is the unique integer remainder of n divided by 2.

Moreover

$$\|y - \mathbf{1}x\|_1 = \|a - \mathbf{1}x\|_1, \qquad x \in \mathbb{R}.$$

Therefore, to prove the lemma, it is enough to show that

$$|a|_1 = \min_{x\in\mathbb{R}} \|a - \mathbf{1}x\|_1 = \sum_{i>(q+r)} a_i - \sum_{i\le q} a_i. \tag{9}$$

Suppose n is even in which case $n = 2q$ and $r = 0$. If $q = 1$, then $\min_{x\in\mathbb{R}}(|a_1 - x| + |a_2 - x|) = a_2 - a_1$ in which case (9) holds. Suppose $q > 1$. For fixed $k \in \{1, 2, \ldots, q-1\}$ and any value of x located in the interval $[a_k, a_{k+1}]$, it must be true that $|a_i - x| = x - a_i$ for $i \le k$ and $|a_i - x| = a_i - x$ for $i \ge k+1$. Since $k < q$, the number of values of i such that $i \ge k+1$ is greater than the number of values of i such that $i \le k$. Thus, for $x \in [a_k, a_{k+1}]$, the sum $\sum_{i=1}^n |a_i - x|$ is a linear polynomial in x and the coefficient of x is negative. Since this is true for all $k \in \{1, 2, \ldots, q-1\}$, the sum $\sum_{i=1}^n |a_i - x|$ is a decreasing function of x on the union of the intervals $[a_k, a_{k+1}]$, $k = 1, 2, \ldots, q-1$; i.e., on $[a_1, a_q]$. By similar reasoning $\sum_{i=1}^n |a_i - x|$ is an increasing function of x on the interval $[a_{q+1}, a_n]$. Meanwhile, for values of $x \in [a_q, a_{q+1}]$, clearly $|a_q - x| = x - a_q$ and $|a_{q+1} - x| = a_{q+1} - x$, so the sum $\sum_{i=1}^n |a_i - x|$ is a constant. But $\sum_{i=1}^n |a_i - x|$ is a continuous function of x. Therefore, $\sum_{i=1}^n |a_i - x|$ is nonincreasing for $x \le a_q$ and nondecreasing for $x \ge a_q$. Therefore, a value of x which minimizes $\sum_{i=1}^n |a_i - x|$ is $x = a_q$. Equation (9) follows at once.

Now suppose n is odd in which case $r = 1$. For fixed $k \in \{1, 2, \ldots, q\}$ and any value of x located in the interval $[a_k, a_{k+1}]$, it must be true that $|a_i - x| = x - a_i$ for $i \le k$ and $|a_i - x| = a_i - x$ for $i \ge k+1$. Since $k \le q$, the number of values of i such that $i \ge k+1$ is greater than the number of values of i such that $i \le k$. Thus, for $x \in [a_k, a_{k+1}]$, the sum $\sum_{i=1}^n |a_i - x|$ is a linear polynomial in x and the coefficient of x is negative. Since this is true for all $k \in \{1, 2, \ldots, q\}$, the sum $\sum_{i=1}^n |a_i - x|$ is a decreasing function of x on the union of the intervals $[a_k, a_{k+1}]$, $k = 1, 2, \ldots, q$; i.e., on $[a_1, a_{q+1}]$. By similar reasoning $\sum_{i=1}^n |a_i - x|$ is an increasing function of x on the interval $[a_{q+1}, a_n]$. Therefore, the unique value of x which minimizes $\sum_{i=1}^n |a_i - x|$ is $x = a_{q+1} = a_{q+r}$. Equation (9) follows at once. ■

Consider now the case when M is a doubly stochastic matrix S. Then, the column sums of S are all equal to 1. This implies that $|S|_1 \le 1$ because $|S|_1 \le \|S\|_1 = 1$. The column sums all equaling one also imply that

$$\sum_{i\in\mathcal{L}_j} s_{ij} + rm_j + \sum_{i\in\mathcal{S}_j} s_{ij} = 1, \qquad j \in \{1, 2, \ldots, n\}$$

where m_j is the median[3] of the n entries in the jth column of S. Therefore

$$|S|_1 = \max_{j\in\{1,2,\ldots,n\}} \left\{ 2\sum_{i\in\mathcal{L}_j} s_{ij} + rm_j - 1 \right\}.$$

This means that S is semicontractive in the one-norm just in case

$$\sum_{i\in\mathcal{L}_j} s_{ij} + \frac{r}{2} m_j < 1, \qquad j \in \{1, 2, \ldots, n\}.$$

We are led to the following result.

Theorem 6: Let q be the unique integer quotient of n divided by 2. Let $S \in \mathbb{R}^{n\times n}$ be a doubly stochastic matrix. Then, $|S| \le 1$. Moreover, S is a semicontraction in the one-norm if and only if the number of nonzero entries in each column of S exceeds q.

Note that the doubly stochastic matrix

$$S = \begin{bmatrix} 0.5 & 0.125 & 0.125 & 0.125 & 0.125 \\ 0.5 & 0.125 & 0.125 & 0.125 & 0.125 \\ 0 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0 & 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix}$$

has a weakly connected graph but is not a semicontraction for $p = 1$. Thus, this particular seminorm is not as useful as we would like for gossiping problems.

It is possible to compare this seminorm with the coefficient of ergodicity. Observe that while the preceding matrix is not a semicontraction it is a scrambling matrix. Thus, for this example, $\tau(S) < |S|_1 = 1$. On the other hand, there are also doubly stochastic matrices which are semicontractions but which are not scrambling matrices. An example of this is the matrix

$$S = \begin{bmatrix} 0.5 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0.5 \\ 0.125 & 0.125 & 0.25 & 0.25 & 0.125 & 0.125 \\ 0.125 & 0.125 & 0.25 & 0.25 & 0.125 & 0.125 \\ 0.125 & 0.125 & 0.25 & 0.25 & 0.125 & 0.125 \end{bmatrix}.$$

Thus, for this example, $|S|_1 < \tau(S) = 1$, which means that there are situations when it may be more advantageous to

[3]The *median* of a finite set of real numbers is the "middle value" of the set. More precisely, suppose that $\mathcal{A}$ is a set of n real numbers which are $a_1 \le a_2 \le \cdots \le a_n$ and let q and r be, respectively, the unique integer quotient and remainder of n divided by 2. If n is odd, the median of $\mathcal{A}$ is a_{q+r}. If n is even, the median of $\mathcal{A}$ is defined to be the average of a_q and a_{q+1}.

use the seminorm $|\cdot|_1$ to compute convergence rates than to appeal to the coefficient of ergodicity.

2) Case $p = \infty$: Note that in this case $|S|_\infty \leq 1$ for any stochastic matrix because $|S|_\infty \leq \|S\|_\infty = 1$. Although not at all obvious, it turns out that $|S|_\infty$ equals the well-known coefficient of ergodicity discussed earlier and defined by (1). This is an immediate consequence of Proposition 5 which is stated below. Unfortunately, the last example in the preceding section shows that there are doubly stochastic matrices with weakly connected graphs which are not scrambling matrices. Thus, this particular seminorm is also not useful for our purposes.

Proposition 5: Let $A \in \mathbb{R}^{n\times n}$ be a nonnegative matrix. Then

$$|A|_\infty = \frac{1}{2}\max_{i,j}\sum_{k=1}^{n}|a_{ik} - a_{jk}|.$$

The proof of Proposition 5 depends on the following lemma.

Lemma 3: Suppose $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$ is a set of $m > 1$ row vectors in $\mathbb{R}^{1\times n}$. Let $d(x, y)$ denote the metric

$$d(x,y) = \sum_{i=1}^{n}|x_i - y_i|, \qquad x, y \in \mathbb{R}^{1\times n}$$

where x_i and y_i are the ith entries of x and y, respectively. Then

$$\min_{c\in\mathbb{R}^{1\times n}}\max_i d(c, a_i) = \frac{1}{2}\max_{j,k} d(a_j, a_k).$$

Proof of Lemma 3: For any j and k, and any row vector $c \in \mathbb{R}^{1\times n}$, $d(a_j, c) + d(c, a_k) \geq d(a_j, a_k)$ because of the triangle inequality. Clearly $\max_i d(c, a_i) \geq (1/2)(d(a_j, c) + d(c, a_k))$, so

$$\max_i d(c, a_i) \geq \frac{1}{2}d(a_j, a_k)$$

for all j, k. In particular, $\max_i d(c, a_i) \geq (1/2)\max_{j,k} d(a_j, a_k)$. Since this is true for all c

$$\min_c \max_i d(c, a_i) \geq \frac{1}{2}\max_{j,k} d(a_j, a_k).$$

To complete the proof, it must be shown that

$$\min_c \max_i d(c, a_i) \leq \frac{1}{2}\max_{j,k} d(a_j, a_k).$$

For this it is enough to show that there exists a closed ball $\mathcal{B}$ in $\mathbb{R}^{1\times n}$ with radius $r = (1/2)\max_{j,k} d(a_j, a_k)$ such that $\mathcal{B} \supset \mathcal{A}$. Towards this end[4] let $\mathcal{B}$ denote the ball

$$\mathcal{B} = \left\{x : x \in \mathbb{R}^{1\times n}, \sum_{i=1}^{n}|x_i| \leq r\right\}.$$

The boundary of $\mathcal{B}$, $\partial\mathcal{B}$, consists of 2^{n-1} pairs of parallel $(n-1)$-hyperplanes. Consider one of these pairs, $\mathcal{H}_1 : x_1 + x_2 + \cdots + x_n = r$ and $\mathcal{H}_2 : x_1 + x_2 + \cdots + x_n = -r$. For any $y \in \mathcal{H}_1$ and $z \in \mathcal{H}_2$

$$\sum_{i=1}^{n}|y_i - z_i| \geq \sum_{i=1}^{n}(y_i - z_i) = 2r$$

which shows that the distance of the two hyperplanes equals $2r$. Thus, there must exist a real number s such that every $a_i \in \mathcal{A}$ lies between or on the two parallel hyperplanes

$$x_1 + x_2 + \cdots + x_n = s + r$$
$$x_1 + x_2 + \cdots + x_n = s - r.$$

Using the same arguments, the other $2^{n-1} - 1$ pairs of parallel hyperplanes also can be shown to exist. Therefore, $\mathcal{A}$ can be contained in a closed set bounded by the 2^{n-1} pairs of parallel $(n-1)$-hyperplanes which is a closed ball with radius $(1/2)\max_{j,k} d(a_j, a_k)$ in $\mathbb{R}^{1\times n}$. ■

3) Case $p = 2$: For the case when $p = 2$ it is also possible to derive an explicit formula for the seminorm $|M|_2$ of a nonnegative matrix $M \in \mathbb{R}^{n\times n}$. Towards this end note that for any $x \in \mathbb{R}^n$, the function

$$g(x, c) = x'(M - \mathbf{1}c)'(M - \mathbf{1}c)x$$
$$= x'M'Mx - 2x'M'\mathbf{1}cx + n(cx)^2$$

[4]We are indebted to Chun-Yi Sun (Department of Mathematics, Yale University) for pointing this out to us.

attains its minimum with respect to c at

$$c = \frac{1}{n}\mathbf{1}'M.$$

This implies that

$$|M|_2 = \|PM\|_2 = \sqrt{\mu\{M'P'PM\}}$$

where $P = I - (1/n)\mathbf{1}\mathbf{1}'$ and, for any symmetric matrix T, $\mu\{T\}$ is the largest eigenvalue of T. We are led to the following result.

Proposition 6: Let $M \in \mathbb{R}^{n\times n}$ be a nonnegative matrix. Then, $|M|_2$ is the largest singular value of the matrix PM where P is the orthogonal projection on the orthogonal complement of the span of $\mathbf{1}$.

Now suppose that M is a doubly stochastic matrix S. Then, $S'S$ is also doubly stochastic and $\mathbf{1}'S = \mathbf{1}'$. The latter and Proposition 6 imply that

$$|S|_2 = \sqrt{\mu\left\{S'S - \frac{1}{n}\mathbf{1}\mathbf{1}'\right\}}.$$

More can be said.

Lemma 4: If S is doubly stochastic, then $\mu\{S'S - (1/n)\mathbf{1}\mathbf{1}'\}$ is the second largest eigenvalue of $S'S$.

Proof of Lemma 4: Since $S'S$ is symmetric it has orthogonal eigenvectors one of which is $\mathbf{1}$. Let $\mathbf{1}, x_2, \ldots, x_n$ be such a set of eigenvectors with eigenvalues $1, \lambda_2, \ldots, \lambda_n$. Then, $S'S\mathbf{1} = \mathbf{1}$ and $S'Sx_i = \lambda_i x_i$, $i \in \{2, 3, \ldots, n\}$. Clearly $(S'S - (1/n)\mathbf{1}\mathbf{1}')\mathbf{1} = 0$ and $(S'S - (1/n)\mathbf{1}\mathbf{1}')x_i = \lambda_i x_i$, $i \in \{2, 3, \ldots, n\}$. Since 1 is the largest eigenvalue of $S'S$ it must therefore be true that the second largest eigenvalue $S'S$ is the largest eigenvalue of $S'S - (1/n)\mathbf{1}\mathbf{1}'$. ■

We summarize as follows.

Theorem 7: For $p = 2$, the seminorm of a doubly stochastic matrix S is the second largest singular value of S.

There is another way to think about what this theorem implies. Prompted by the work in [8] and [23], suppose one wants to measure in the sense of a 2-norm $\|\cdot\|$, how much closer an n-vector x gets to the average vector $z = (1/n)\mathbf{1}\mathbf{1}'x$ when it is multiplied by a doubly stochastic matrix S. In other words how does the norm $\|Sx - z\|$ compare with $\|x - z\|$? To address this question, note first that $x - z \in \mathcal{O}$ where $\mathcal{O}$ is the orthogonal complement of the span of $\mathbf{1}$. Note next that

$$\|Sx - z\|^2 = \|S(x - z)\|^2 \leq \left(\sup_{y\in\mathcal{O}} \frac{y'S'Sy}{y'y}\right)\|x - z\|^2.$$

But $\sup_{y\in\mathcal{O}} y'S'Sy/y'y$ is the second largest eigenvalue of $S'S$ which in turn is the square of the second largest singular value of S. In other words, $\|Sx - z\| \leq |S|_2\|x - z\|$. Thus, Sx is always as close to the average vector z as x is and is even closer if $|S|_2$ is a contraction.

In the light of Theorem 7, we are now in a position to characterize in graph-theoretic terms those doubly stochastic matrices with positive diagonal entries which are semicontractions for $p = 2$.

Theorem 8: Let S be a doubly stochastic matrix with positive diagonal entries. Then, $|S|_2 \leq 1$. Moreover, S is a semicontraction in the 2-norm if and only if the graph of S is weakly connected.

To prove this theorem we need several concepts and results. Let $\mathbb{G}$ denote a directed graph and write $\mathbb{G}'$ for that graph which results when the arcs in $\mathbb{G}$ are reversed; i.e., the *dual graph*. Call a graph *symmetric* if it is equal to its dual. Note that in the case of a symmetric graph, the three properties of being rooted, strongly connected, and weakly connected are equivalent. Note also that if $\mathbb{G}$ is the graph of a nonnegative matrix M with positive diagonal entries, then $\mathbb{G}'$ is the graph of M' and $\mathbb{G}' \circ \mathbb{G}$ is the graph of $M'M$.

Lemma 5: A directed graph $\mathbb{G}$ with self-arcs at all vertices is weakly connected if and only if $\mathbb{G}' \circ \mathbb{G}$ is strongly connected.

Proof of Lemma 5: Since $\mathbb{G}$ has self-arcs at all vertices so does $\mathbb{G}'$. This implies that the arc set of $\mathbb{G}' \circ \mathbb{G}$ contains the arc sets of $\mathbb{G}$ and $\mathbb{G}'$. Thus, for any undirected path in $\mathbb{G}$ between vertices i and j, there must be a corresponding directed path in $\mathbb{G}' \circ \mathbb{G}$ between the same two vertices. Thus, if $\mathbb{G}$ is weakly connected, $\mathbb{G}' \circ \mathbb{G}$ must be strongly connected.

Now suppose that (i, j) is an arc in $\mathbb{G}' \circ \mathbb{G}$. Then, because of the definition of composition, there must be a vertex k such that (i, k) is an arc in $\mathbb{G}$ and (k, j) is an arc in $\mathbb{G}'$. This implies that (i, k) and (j, k) are arcs in $\mathbb{G}$. Thus, $\mathbb{G}$ has an undirected path from i to j. Now suppose that $(i, v_1), (v_1, v_2), \ldots, (v_q, j)$ is a directed path in $\mathbb{G}' \circ \mathbb{G}$ between i and j. Between each pair of successive vertices along this path there must therefore be an undirected path in $\mathbb{G}$. Thus, there must be an undirected path in $\mathbb{G}$ between i and j. It follows that if $\mathbb{G}' \circ \mathbb{G}$ is strongly connected, then $\mathbb{G}$ is weakly connected. ■

Lemma 6: Let T be a stochastic matrix with positive diagonal entries. If T has a strongly connected graph, then the magnitude of its second largest eigenvalue is less than 1. If, on the other hand, the magnitude of the second largest eigenvalue of T is less than 1, then the graph of T is weakly connected.

Proof of Lemma 6: Suppose that the graph of T is strongly connected. Then, via Theorem 6.2.24 of [27], T is irreducible. Thus, there is an integer k such that $(I+T)^k > 0$. Since T has positive diagonal entries, this implies that $T^k > 0$. Therefore, T is primitive [27]. Thus, by the Perron–Frobenius theorem [26], T can have only one eigenvalue of maximum modulus. Since the spectral radius of T is 1 and 1 is an eigenvalue, the magnitude of the second largest eigenvalue of T must be less than 1.

To prove the converse, suppose that T is a stochastic matrix whose second largest eigenvalue in magnitude is less than 1. Then

$$\lim_{i\to\infty} T^i = \mathbf{1}c$$

for some row vector c. Suppose that the graph of T is not weakly connected. Therefore, if q denotes the number of weakly connected components of the graph, then $q > 1$. This implies that $T = P'DP$ for some permutation matrix P and block diagonal matrix D with q blocks. Since $D = PTP'$, D is also stochastic. Thus, each of its q diagonal blocks is stochastic. Since T^i converges to $\mathbf{1}c$, D^i must converge to a matrix of the form $\mathbf{1}\bar{c}$. But this is clearly impossible because $\mathbf{1}\bar{c}$ cannot have $q > 1$ diagonal blocks. ■

Proof of Theorem 8: Let S be a doubly stochastic matrix with positive diagonal entries. Then, 1 is the largest singular value of S because $S'S$ is doubly stochastic. From this and theorem 7 it follows that $|S|_2 \leq 1$.

Suppose S is a semicontraction. Then, in view of Theorem 7, the second largest eigenvalue of $S'S$ is less than 1. Thus, by Lemma 6, the graph of $S'S$ is weakly connected. But $S'S$ is symmetric so its graph must be strongly connected. Therefore, by Lemma 5, the graph of S is weakly connected.

Now suppose that the graph of S is weakly connected. Then, the graph of $S'S$ is strongly connected because of Lemma 5. Thus, by Lemma 6, the magnitude of the second largest eigenvalue of $S'S$ is less than 1. From this and Theorem 7 it follows that S is a semicontraction. ■

Proof of Theorem 5: Let $\mathcal{M}$ be any compact subset of $\mathcal{W}$. In view of Theorem 8, each matrix in $\mathcal{M}$ is a semicontraction in the 2-norm. From this and Proposition 3, it follows that $\mathcal{M}$ is convergable.

Now suppose that $\mathcal{M}$ is convergable and let S be a matrix in $\mathcal{M}$. Then, S^i converges to a matrix of the form $\mathbf{1}c$ as $i \to \infty$. This means that the second largest eigenvalue of S must be less than 1 in magnitude. Thus, by Lemma 6, S must have a weakly connected graph. ■

The importance of Theorem 8 lies in the fact that the matrices in every convergable set of doubly stochastic matrices are contractions in the 2-norm. In view of Proposition 3, this enables one to immediately compute a rate of convergence for any infinite product of matrices from any given convergable set. The coefficient of ergodicity mentioned earlier does not have this property. If it did, then every doubly stochastic matrix with a weakly connected graph would have to be a scrambling matrix. The following counterexample shows that this is not the case:

$$S = \begin{bmatrix} 0.5 & 0.25 & 0 & 0 & 0 & 0.25 \\ 0.25 & 0.5 & 0 & 0 & 0 & 0.25 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.25 & 0 & 0.25 \\ 0 & 0 & 0 & 0 & 0.875 & 0.125 \\ 0.25 & 0.25 & 0 & 0.25 & 0.125 & 0.125 \end{bmatrix}.$$

In particular, S is a doubly stochastic matrix with a weakly connected graph but it is not a scrambling matrix.

D. Contraction Coefficient

By the *contraction coefficient* of a gossip matrix M is meant the seminorm $|M|_2$. The main result we want to prove is as follows.

Theorem 9: A gossip matrix M is complete if and only if its contraction coefficient is less than one.

Before turning to a proof of this theorem, let us consider its consequences. As in the hypothesis of Theorem 1, let $M(1), M(2), M(3), \ldots$ denote the gossiping matrices corresponding to an infinite sequence of single gossips which is repetitively complete with period T. Our aim is to explain how fast the matrix product $M(t)M(t-1)\cdots M(1)$ converges to $(1/n)\mathbf{1}\mathbf{1}'$, and in so doing provide a proof of Theorem 1. Towards this end note first that each $M(t) \in \mathcal{P}$, the set of all $n \times n$ single-gossip matrices. Let $\mathcal{C}$ denote the set of all complete gossip matrices which are products of exactly T single-gossip matrices from $\mathcal{P}$. Note that $\mathcal{C}$ is a compact set because $\mathcal{P}$ is. For each $S \in \mathcal{C}$, let $|S|_2$ denote the contraction coefficient of S. Then, in view of Theorem 9, $|S|_2 < 1$, $S \in \mathcal{C}$. Therefore, the nonnegative number

$$\mu = \max_{\mathcal{C}} |S|_2$$

is less than one. Next observe that since $M(1), M(2), M(3), \ldots$ corresponds to a repetitively complete gossip sequence, each matrix $N_i = M(iT)M(iT-1)((i-1)T+1)$, $i \geq 1$, must be in $\mathcal{C}$; thus, $|N_i|_2 \leq \mu$, $i \geq 1$. It follows from Proposition 3 that as $i \to \infty$, the matrix product $N_iN_{i-1}\cdots N_1$ converges to $(1/n)\mathbf{1}\mathbf{1}'$ as fast at μ^i converges to zero. Thus, if we define $\lambda = \mu^{1/T}$, then as $t \to \infty$, the

matrix product $M(t)M(t-1)\cdots M(1)$ also converges to $(1/n)\mathbf{1}\mathbf{1}'$, in this case as fast as λ^t converges to zero. The final assertion of Theorem 1 follows at once.

We now turn to the proof of Theorem 9. For this we need several preliminary results.

Lemma 7: Let $\mathbb{G}$ and $\mathbb{H}$ be directed graphs on the same n vertices. Suppose that both graphs have self-arcs at all vertices. If there is an undirected path from i to j in $\mathbb{H}\circ\mathbb{G}$, then there is an undirected path from i to j in the union of $\mathbb{H}$ and $\mathbb{G}$.

Proof of Lemma 7: First suppose that there is an undirected path of length one between vertices i and j in $\mathbb{H}\circ\mathbb{G}$. Then, either (i,j) or (j,i) must be an arc in $\mathbb{H}\circ\mathbb{G}$. Without loss of generality suppose that (i,j) is an arc in $\mathbb{H}\circ\mathbb{G}$. Then, because of the definition of composition, there must be a vertex k such that (i,k) is an arc in $\mathbb{G}$ and (k,j) is an arc in $\mathbb{H}$. This implies that (i,k) and (k,j) are arcs in $\mathbb{H}\cup\mathbb{G}$. Thus, $\mathbb{H}\cup\mathbb{G}$ has a directed path from i to j. Therefore, $\mathbb{H}\cup\mathbb{G}$ has an undirected path from i to j. Now suppose that $(i,v_1),(v_1,v_2),\ldots,(v_q,j)$ is an undirected path in $\mathbb{H}\circ\mathbb{G}$ between i and j. Between each pair of successive vertices along this path there must therefore be an undirected arc in $\mathbb{H}\circ\mathbb{G}$. Therefore, between each pair of successive vertices along this path there must be an undirected arc in $\mathbb{H}\cup\mathbb{G}$. Thus, there must be an undirected path in $\mathbb{H}\cup\mathbb{G}$ between i and j. ■

It is obvious that the preceding lemma extends from two graphs to a finite set of directed graphs. In the following, we appeal to this extension without special mention. The next result characterizes completeness of a gossip matrix in terms of a property of its graph.

Lemma 8: A gossip matrix is complete if and only if its graph is weakly connected.

Proof of Lemma 8: Since M is a gossip matrix, there exist single-gossip matrices $P_1,P_2,\ldots,P_m$ such that $M=P_mP_{m-1}\cdots P_1$. Write (j_i,k_i) for the edge in $\mathbb{N}$ associated with P_i and let $\mathbb{G}$ denote the spanning subgraph of $\mathbb{N}$ whose edges are $(j_1,k_1),(j_2,k_2),\ldots,(j_m,k_m)$. Since each P_i is a single-gossip matrix, its graph $\gamma(P_i)$ must contain arcs from j_i to k_i and from k_i to j_i. This means that the union of the $\gamma(P_i)$, written $\mathbb{U}$, will be weakly connected if and only if $\mathbb{G}$ is a connected graph. Therefore, weak connectivity of $\mathbb{U}$ is equivalent to M being complete.

In general, the composition of two directed graphs contains the union of the two graphs whenever the two graphs in question have self-arcs at all vertices. This means that $\mathbb{U}$ must be a subgraph of $\gamma(M)$ because each vertex of each $\gamma(P_i)$ has a self-arc. But it has just been shown that completeness of M implies weak connectivity of $\mathbb{U}$. It follows that completeness of M must also imply weak connectivity of $\gamma(M)$.

Suppose next that $\gamma(M)$ is weakly connected. Then, $\mathbb{U}$ must be weakly connected because of Lemma 7. Therefore, M must be complete. ■

Proof of Theorem 9: Since M is a product of matrices with positive diagonal entries, M has positive diagonal entries. By Lemma 8, M is complete if and only if $\gamma(M)$ is weakly connected. By Theorem 8, any doubly stochastic matrix with positive diagonal entries has a contraction coefficient less than one just in case the graph of the matrix is weakly connected. Therefore, M is complete if and only if it has a contraction coefficient less than one. ■

As it stands the proof of Theorem 1 rests on the assumption that $\gamma(M)$ has a weakly connected graph. Nedić *et al.* [23] derives a result very similar to Theorem 1 under what at first glance appears to be a more restrictive assumption, namely that $\gamma(M)$ has a *strongly* connected graph. But because doubly stochastic matrices constitute a special class of stochastic matrices, the assumption of strong connectivity turns out to be not restrictive at all. Here is why.

Lemma 9: The graph $\mathbb{G}$ of a doubly stochastic matrix D is strongly connected if and only if it is weakly connected.[5]

The proof of Lemma 9 which follows is based on ideas from [26] and [34]. Let $\mathbb{G}$ be a directed graph with vertex set $\mathcal{V}=\{1,2,\ldots,n\}$. Call a vertex j *reachable* from i if either $j=i$ or if there is a directed path from i to j. Call a vertex i *essential* if i is reachable from all vertices which are reachable from i.

Lemma 10: Every directed graph has at least one essential vertex.

Proof of Lemma 10: Suppose that $\mathbb{G}$ has $n>0$ vertices. If $\mathbb{G}$ has an isolated vertex i, then i is essential. Consider next the case when $\mathbb{G}$ has no isolated vertices in which case $n>1$. Suppose that $\mathbb{G}$ has no essential vertices. Then, for each vertex i there must be a vertex j which is reachable from i but from which i is not reachable. It follows that it is possible to construct a sequence of vertices $i_1,i_2,\ldots,i_m$ of any length $m>1$ such that i_{j+1} is reachable from i_j, $j\in\{1,2,\ldots,m-1\}$ but not conversely. But for $m>n$ at least one vertex must appear in the list twice, say in positions j and $k>j$. This implies that vertex $j+1$ is reachable from j and conversely which is a contradiction. Therefore, $\mathbb{G}$ must have an essential vertex. ■

To proceed, let us say that vertices i and j are *mutually reachable* if each is reachable from the other. Mutual reachability is clearly an equivalence relation on $\mathcal{V}$ which partitions $\mathcal{V}$ into the disjoint union of a finite number of equivalence classes. Note that if i is an essential vertex of $\mathbb{G}$, then every vertex in the equivalence class of i is also essential. Thus, every directed graph possesses at least one mutually reachable equivalence class whose members are all essential.

[5]It is clear that strong connectivity of $\mathbb{G}$ implies weak connectivity of $\mathbb{G}$. The converse was conjectured by John Tsitsiklis in a private communication.

Proof of Lemma 9: Strong connectivity clearly implies weak connectivity. We prove the converse. Suppose $\mathbb{G}$ is weakly connected. In view of the proceeding, $\mathbb{G}$ has at least one mutually reachable equivalence class $\mathcal{E}$ whose members are all essential. If $\mathcal{E} = \mathcal{V}$, then $\mathbb{G}$ is obviously strongly connected. Thus, to prove the lemma, it is enough to show that $\mathcal{E} = \mathcal{V}$. Suppose the contrary, namely that $\mathcal{E} = \{i_1, i_2, \ldots, i_m\}$ is a strictly proper subset of $\mathcal{V}$. Let π be any permutation map for which $\pi(i_j) = j$, $j \in \{1, 2, \ldots, m\}$, and let P be the corresponding permutation matrix. Then, clearly

$$P'DP = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

and $P'DP$ is doubly stochastic. Since $P'DP$ is doubly stochastic, the column sums of A must all equal one as must the row sums of the submatrix $[A\ B]$. But the transformation $D \longmapsto P'DP$ corresponds to a relabeling of the vertices of $\mathbb{G}$, so the graph of $P'DP$ must also be weakly connected. This means that B cannot be the zero matrix. Therefore, the sum of the row sums of A must be less than m. But this contradicts the fact that the sum of the column sums of A equals m. Therefore, $\mathcal{E} = \mathcal{V}$. ■

We are now in a position to prove Theorem 2.

Proof of Theorem 2: Suppose that M is a complete gossip matrix. In view of Lemmas 8 and 9, $\gamma(M)$ is a strongly connected graph. Thus, by Lemma 6, the magnitude of the second largest eigenvalue of M is less than 1.

To prove the converse, now suppose that M is a gossip matrix whose second largest eigenvalue in magnitude is less than one. By Lemma 6 the graph of M is therefore weakly connected. Thus, by Lemma 8, M is complete. ■

V. CONCLUDING REMARKS

Let $\mathbb{N}$ be a given neighbor graph. Recall that a complete gossip sequence is *minimal* if there is no shorter sequence of gossips which is compete. It is easy to see that a complete gossip sequence will be minimal if and only if the gossip graph it induces is a minimal spanning tree in $\mathbb{N}$. For a given neighbor graph there can be many minimal spanning trees and consequently many minimally complete gossip sequences. Moreover, there can be differing second largest singular values and different second largest eigenvalues (in magnitude) for the different doubly stochastic matrices associated with different complete minimal sequences. A useful exercise then would be to determine those complete minimal sequences whose associated singular values or second largest eigenvalues (in magnitude) are as small as possible.

One of the problems with the idea of gossiping, which apparently is not widely appreciated, is that it is difficult to devise provably correct gossiping protocols which are guaranteed to avoid deadlocks without making restrictive assumptions. The research in this paper and in [12] and [13] contributes to our understanding of this issue and how to deal with it.

REFERENCES

[1] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Dept. Electr. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, 1984.

[2] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Autom. Control*, vol. AC-31, no. 9, pp. 803–812, Sep. 1986.

[3] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988–1001, Jun. 2003.

[4] R. Olfati-Saber and R. M. Murray, "Consensus seeking in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[5] V. D. Blondel, J. M. Hendrichkx, A. Olshevsky, and J. N. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," in *Proc. 44th IEEE Conf. Decision Control*, 2005, pp. 2996–3000.

[6] L. Moreau, "Stability of multi-agent systems with time-dependent communication links," *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 169–182, Feb. 2005.

[7] W. Ren and R. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, May 2005.

[8] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65–78, Jan. 2004.

[9] A. Kashyap, T. Basar, and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192–1203, 2007.

[10] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. 4th Int. Conf. Inf. Process. Sensor Netw.*, 2005, pp. 63–70.

[11] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Gossip algorithms: Design, analysis and applications," in *Proc. 24th IEEE Int. Conf. Comput. Commun.*, 2005, pp. 1653–1664.

[12] M. Mehyar, D. Spanos, J. Pongsajapan, S. H. Low, and R. M. Murray, "Asynchronous distributed averaging on communication networks," *IEEE/ACM Trans. Netw.*, vol. 15, no. 3, pp. 512–520, Jun. 2007.

[13] A. Olshevsky, "Efficient information aggregation strategies for distributed control and signal processing," Ph.D. dissertation, Dept. Electr. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, 2010.

[14] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2508–2530, Jun. 2006.

[15] B. D. O. Anderson, C. Yu, and A. S. Morse, "Convergence of periodic gossiping algorithms," in *Perspectives in Mathematical System Theory, Control, and Signal Processing*, J. C. Willems, S. Hara, Y. Ohta, and H. Fujioka, Eds. New York: Springer-Verlag, 2010, pp. 127–138.

[16] M. Cao, D. A. Spielman, and A. S. Morse, "A lower bound on convergence of a distributed network consensus algorithm," in *Proc. 44th IEEE Conf. Decision Control*, 2005, pp. 2356–2361.

[17] A. Olshevsky and J. N. Tsitsiklis, "Convergence rates in distributed consensus and averaging," in *Proc. 45th IEEE Conf. Decision Control*, 2006, pp. 3387–3392.

[18] M. Cao, D. A. Spielman, and E. M. Yeh, "Accelerated gossip algorithms for distributed computation," in *Proc. 44th Annu. Allerton Conf. Commun. Control Comput.*, 2006, pp. 952–959.

[19] J. Liu, B. D. O. Anderson, M. Cao, and A. S. Morse, "Analysis of accelerated gossip algorithms," in *Proc. 48th IEEE Conf. Decision Control*, 2009, pp. 871–876.

[20] S. Mou, C. Yu, B. D. O. Anderson, and A. S. Morse, "Deterministic gossiping with a periodic protocol," in *Proc. 49th IEEE Conf. Decision Control*, 2010, pp. 5787–5791.

[21] F. He, A. S. Morse, J. Liu, and S. Mou, "Periodic gossiping," in *Proc. 18th Int. Fed. Autom. Control World Congr.*, 2011, to be published.

[22] A. Olshevsky and J. N. Tsitsiklis, "Convergence speed in distributed consensus and averaging," *SIAM J. Control Optim.*, vol. 48, no. 1, pp. 33–55, 2009.

[23] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "On distributed averaging algorithms and quantization effects," *IEEE Trans. Autom. Control*, vol. 54, no. 11, pp. 2506–2517, Nov. 2009.

[24] J. Liu, A. S. Morse, B. D. O. Anderson, and C. Yu, "The contraction coefficient of a complete gossip sequence," in *Three Decades of Progress in Control Sciences*, X. Hu, U. Jonsson, B. Wahlberg, and B. Ghosh, Eds. New York: Springer-Verlag, 2010, pp. 275–290.

[25] C. Yu, B. D. O. Anderson, S. Mou, J. Liu, F. He, and A. S. Morse, "Gossiping periodically," *IEEE Trans. Autom. Control*, 2011.

[26] E. Seneta, *Non-Negative Matrices and Markov Chains*. New York: Springer-Verlag, 2006.

[27] R. C. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.

[28] R. Diestel, *Graph Theory*. Berlin, Germany: Springer-Verlag, 2005.

[29] V. G. Vizing, "On an estimate of the chromatic class of a *p*-graph," *Diskret Analiz*, vol. 3, pp. 25–30, 1964.

[30] R. L. Brooks, "On the colouring of nodes of a network," *Math. Proc. Cambridge Philosoph. Soc.*, vol. 37, no. 2, pp. 194–197, 1941.

[31] J. Liu, S. Mou, A. S. Morse, B. D. O. Anderson, and C. Yu, "Request-based gossiping," in *Proc. 50th IEEE Conf. Decision Control*, 2011, to appear.

[32] J. Wolfowitz, "Products of indecomposable, aperiodic, stochastic matrices," *Proc. Amer. Math. Soc.*, vol. 15, no. 5, pp. 733–737, 1963.

[33] M. Cao, A. S. Morse, and B. D. O. Anderson, "Reaching a consensus in a dynamically changing environment: A graphical approach," *SIAM J. Control Optim.*, vol. 47, no. 2, pp. 575–600, 2008.

[34] R. G. Gallager, *Discrete Stochastic Processes*. Boston, MA: Kluwer, 1996.

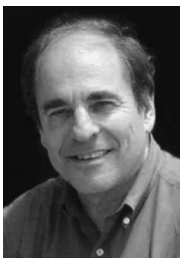
ABOUT THE AUTHORS

Ji Liu (Student Member) [redacted] He received the B.S. degree in information engineering from Shanghai Jiao Tong University, Shanghai, China, in 2006 and the M.S. degree in electrical engineering from Yale University, New Haven, CT, in 2007, where he is currently working towards the Ph.D. degree in electrical engineering.

His research interests are in the areas of multiagent systems and cooperative control.

Shaoshuai Mou (Student Member) [redacted] He received the B.Eng. and M.Eng. degrees in aerospace engineering from Harbin Institute of Technology, Harbin, China, in 2006 and 2008, respectively, and the M.S. degree in electrical engineering from Yale University, New Haven, CT, in 2010, where he is currently working towards the Ph.D. degree in electrical engineeering under the supervision of Prof. Morse.

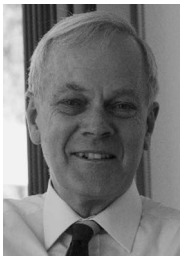
He worked as a Research Associate in the University of Hong Kong, Hong Kong, in 2008. His major research interests include distributed averaging in sensor networks and formation control of autonomous agents.



A. Stephen Morse (Life Fellow, IEEE) received the B.S. degree from Cornell University, Ithaca, NY, the M.S. degree from the University of Arizona, Tucson, and the Ph.D. degree from Purdue University, West Lafayette, IN in 1967, all in electrical engineering .

Following three years at the Office of Control Theory and Application (OCTA), NASA Electronics Research Center, Cambridge, MA, he joined the Yale University, New Haven, CT, faculty, where he is currently the Dudley Professor of Engineering. His main interest is in network synthesis, optimal control, multivariable control, adaptive control, urban transportation, vision-based control, hybrid and nonlinear systems, and, most recently, coordination and control of large grouping of mobile autonomous agents.

Dr. Morse received the George S. Axelby Outstanding Paper Award from the IEEE Control Systems Society in 1993 and 2005, the American Automatic Control Council's Best Paper Award (twice), and the IEEE Technical Field Award for Control Systems in 1999. He is a Distinguished Lecturer of the IEEE Control Systems Society and a member of the National Academy of Engineering and the Connecticut Academy of Science and Engineering.



Brian D. O. Anderson (Life Fellow, IEEE) was born in Sydney, Australia. He was educated in mathematics and electrical engineering at Sydney University, Sydney, N.S.W., Australia, and received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1966.

He is a Distinguished Professor at the Australian National University, Canberra, A.C.T., Australia, and Distinguished Researcher at the National ICT Australia (NICTA). His current research interests are in distributed control, sensor networks, and econometric modeling.

Dr. Anderson is a Fellow of the Australian Academy of Science, the Australian Academy of Technological Sciences and Engineering, the Royal Society, and a foreign associate of the National Academy of Engineering. His awards include the IEEE Control Systems Award of 1997, the 2001 IEEE James H. Mulligan, Jr. Education Medal, and the Bode Prize of the IEEE Control System Society in 1992, as well as several IEEE best paper prizes.

Changbin (Brad) Yu (Senior Member, IEEE) received the B.Eng. degree with first class honors in computer engineering from Nanyang Technological University, Singapore, in 2004 and the Ph.D. degree in information engineering from the Australian National University, Canberra, A.C.T., Australia, in 2008.

Since then he has been on an academic staff at the Australian National University and Adjunct Researcher at the National ICT Australia (NICTA). His current research interests include control of autonomous formations, multiagent systems, mobile sensor networks, human–robot interaction and graph theory.

Dr. Yu had won a competitive Australian Postdoctoral Fellowship (APD) in 2007 and a prestigious Queen Elizabeth II Fellowship (QEII) in 2010. He was also a recipient of Australian Government Edeavour Asia Award (2005), Inaugural Westpac Australian Chinese Students New Age Award (2006), Chinese Government Outstanding Overseas Students Award (2006), *Asian Journal of Control* Best Paper Award (2006–2009), etc.

Analysis of Target Velocity and Position Estimation via Doppler-Shift Measurements

Iman Shames, Adrian N. Bishop, Matthew Smith and Brian D.O. Anderson

***Abstract*— This paper outlines the problem of doppler-based target position and velocity estimation using a sensor network. The minimum number of doppler shift measurements at distinct generic sensor positions to have a finite number of solutions, and later, a unique solution for the unknown target position and velocity is stated analytically, for the case when no measurement noise is present. Furthermore, we study the same problem where not only doppler shift measurements are collected, but also other types of measurements are available, e.g. bearing or distance to the target from each of the sensors. Subsequently, allowing nonzero measurement noise, we present an optimization method to estimate the position and the velocity of the target. An illustrative example is presented to show the validity of the analysis and the performance of the estimation method proposed. Some concluding remarks and future work directions are presented in the end.**

***Index Terms*— Doppler Measurements, Localization, Motion Estimation, Polynomial Optimization**

I. Introduction

Using doppler-shifts for position and velocity estimation has a long history; see e.g. [1]–[8]. Recently, the doppler effect has gained a renewed interest and it has been implemented for cooperative positioning in vehicular networks [9].

In this paper, we consider a scenario with n nodes with both transmitting and sensing capabilities, that are called sensors for the rest of this paper. The target has an unknown position and velocity $\mathbf{x} = [\mathbf{p}^\top \ \mathbf{v}^\top]^\top \in \mathbb{R}^4$. The position of the non-collocated sensors is given by $\mathbf{s}_i = [s_{i,1} \ s_{i,2}]^\top \in \mathbb{R}^2$, $\forall i \in \{1, \dots, n\}$.

The measured doppler-shift is $\hat{f}_i$ at the i^{th} sensor and is caused by a target reflection due to a signal generated earlier by the same sensor. This frequency shift can be approximated by

$$\hat{f}_i = f_i + w_i \tag{1a}$$

$$= 2\frac{f_{c,i}}{c}\left(\frac{(\mathbf{p}-\mathbf{s}_i)^\top}{\|\mathbf{p}-\mathbf{s}_i\|}\right)\mathbf{v} + w_i \tag{1b}$$

where c is the speed of light (or signal propagation) and $\|\cdot\|$ is the standard Euclidean vector norm and $f_{c,i}$ is the carrier frequency employed by this sensor. Finally, w_i is a zero mean Gaussian random variable with known variance σ_i^2. Note here that the localization is to be achieved instantaneously; we are not envisaging collecting information from agents at a number of successive instants of time and using them to infer position at a single instant of time (The connection with filtering methods is explored further below). There are many studies in the literature that try to solve a similar problem via collecting measurements over a time interval and feeding them into an estimator. For example, in [5] the problem of localization of a single aircraft using doppler measurements is studied. Other similar approaches can be found in [4], [6]–[8]. The analysis carried out in this paper along with the optimization method proposed can be considered as constituting a batch processing method for instantaneous estimation of the target location and velocity. This in turn is used to initialize and improve the updates of any other implemented filter which tracks the target position as the target moves in the environment. For example, Kalman-based filters are prone to errors when the target's motion model deviates significantly from the actual target motion. Our analysis can guard against such behavior and re-initialize the filter, e.g. see [10] and references therein.

The main contribution of this paper is that the minimum number of doppler shift measurements required to have a finite number of solutions for the unknown $\mathbf{x}$ is algebraically derived. In some scenarios, a separate piece of knowledge about the target will allow disambiguation. Later, it is extended to the case where having a unique solution is required. Moreover, the scenarios where different types of measurements, e.g. direction-of-arrival, or distance, are available in addition to doppler shift measurements are considered. The aforementioned conclusions assume zero measurement noise; following from this, an optimization method based on polynomial optimization methods is introduced to calculate the velocity and the position of a target where noisy doppler shift measurements are available.

The remaining sections of this paper are organized as follows. In the next section the main problem of interest in considered. In Section III the case where different types of measurement in addition to doppler shift measurement are available to the sensors is considered. A method based on polynomial optimization to estimate the position and the velocity of the target where doppler shift measurements are contaminated by noise is presented in Section IV. An illustrative example presenting the performance of the proposed optimization method is given in Section V. Concluding remarks and future directions come in Section VI.

I. Shames is with the ACCESS Linnaeus Centre, Royal Institute of Technology (KTH) in Stockholm, Sweden. A.N. Bishop and B.D.O. Anderson are with NICTA, Canberra Research Lab and the Australian National University (ANU). M. Smith is with CEA Technologies. This work was supported by the Swedish Research Council (VR) and Knut and Alice Wallenberg Foundation, by USAF-AOARD-10-4102, and by NICTA which is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

II. Required Minimum Number of Doppler Shift Measurements

Initially, it is assumed $f_{c,i}$ is the same for $i = 1, \ldots, n$ and that the measurements are noiseless, i.e. $w_i = 0$.

$$\begin{aligned} \hat{f}_i &= f_i \\ &= 2\frac{\mathbf{v}^\top(\mathbf{p} - \mathbf{s}_i)}{\|\mathbf{p} - \mathbf{s}_i\|}\frac{f_c}{c}. \end{aligned} \tag{2}$$

Normalizing so that $2\frac{f_c}{c} \triangleq 1$, we obtain

$$f_i = \frac{\mathbf{v}^\top(\mathbf{p} - \mathbf{s}_i)}{\|\mathbf{p} - \mathbf{s}_i\|} \tag{3}$$

Now we are ready to pose the problem of interest in this section.

Problem 1. *Consider n stationary sensors at $\mathbf{s}_i \in \mathbb{R}^2$ capable of collecting noiseless doppler shift measurements from a target at position $\mathbf{p}$ moving with a nonzero velocity $\mathbf{v}$ of the form* (2)*.*

1) *What is the minimum value for n such that there is a finite number of solutions for $\mathbf{x}$?*
2) *What is the minimum value for n such that there is a unique solution for $\mathbf{x}$?*

We limit out analysis to the case where the nodes and the target are in $\mathbb{R}^2$, however, note that the analysis for the case that they are in $\mathbb{R}^3$ is much the same.

The answer to the first question posed in Problem 1 is formally presented in the following proposition for the case where $n = 4$ doppler shift measurements are available. This proposition states that with $n = 4$ measurements we have a finite number of solutions and one might be able to disambiguate the solutions using other measurements, e.g. range, or bearing measurements. Moreover, disambiguation may also be made possible due to prior measurements, or to a priori knowledge about the geographic constraints on targets in the area of interest. This particularly is important for the cases where the possible solutions are widely separated, e.g., see Fig. 1.

Proposition 1. *For $n = 4$ doppler measurements as described by* (2) *and generic positions of the sensors there is a finite number of solutions for the unknown $\mathbf{x}$.*

Proof: Denote the noiseless mapping from the agent position and velocity, $\mathbf{x} = [\mathbf{p}^\top \ \mathbf{v}^\top]^\top$ (a vector in $\mathbb{R}^4$) to measurements (another vector in $\mathbb{R}^n$) by F, where n is the number of sensors. More specifically $F(\mathbf{x}) = [f_1 \ f_2 \ f_3 \ f_4]^\top$. Denote J_F to be the Jacobian of F:

$$J_F = \begin{bmatrix} \frac{\partial f_1}{\partial p_1} & \frac{\partial f_1}{\partial p_2} & \frac{\partial f_1}{\partial v_1} & \frac{\partial f_1}{\partial v_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_4}{\partial p_1} & \frac{\partial f_4}{\partial p_2} & \frac{\partial f_4}{\partial v_1} & \frac{\partial f_4}{\partial v_2} \end{bmatrix} \tag{4}$$

where

$$\frac{\partial f_i}{\partial p_1} = \frac{-(p_2 - s_{i,2})(s_{i,2}v_1 - s_{i,1}v_2 + p_1v_2 - p_2v_1)}{\|\mathbf{s}_i - \mathbf{p}\|^3} \tag{5}$$

$$\frac{\partial f_i}{\partial p_2} = \frac{(p_1 - s_{i,1})(s_{i,2}v_1 - s_{i,1}v_2 + p_1v_2 - p_2v_1)}{\|\mathbf{s}_i - \mathbf{p}\|^3} \tag{6}$$

$$\frac{\partial f_i}{\partial v_1} = \frac{(p_1 - s_{i,1})}{\|\mathbf{s}_i - \mathbf{p}\|} \tag{7}$$

$$\frac{\partial f_i}{\partial v_2} = \frac{(p_2 - s_{i,1})}{\|\mathbf{s}_i - \mathbf{p}\|} \tag{8}$$

It is easy to check that for generic values of $\mathbf{x}$, $\mathbf{s}_i$, $i = 1, \cdots, 4$, J_F is not singular. Moreover, we know that the set of solutions to (2) form an algebraic variety. The reason for this is that while (2) is not a polynomial equation, it is easy to see that its zeros are also the zeros of the following polynomial equation.

$$f_i^2\|\mathbf{p} - \mathbf{s}_i\|^2 - (\mathbf{v}^\top(\mathbf{p} - \mathbf{s}_i))^2 = 0 \tag{9}$$

The set of the solutions to the equations described by (9) is known to have at least one member; that is the solution corresponding to the physical setup. The nonsingularity of the Jacobian implies that for generic values for measurements we do not have a continuous set of solutions. As a result, the variety is a zero-dimensional variety. This means that there is a finite number of solutions for the localization problem using doppler measurements where $n \geq 4$. □

Now we briefly consider the case where the Jacobian matrix J_F is singular. This corresponds to those sensor and target geometries where there is an infinite number of solutions for the unknown $\mathbf{p}$ and $\mathbf{v}$. We call these geometries as *bad geometries*. The following proposition characterizes one of these bad geometries.

Proposition 2. *The Jacobian matrix J_F* (4) *is singular if the sensors $\mathbf{s}_1, \cdots, \mathbf{s}_4$ and the target $\mathbf{p}$ are collinear.*

Proof: To prove this proposition it is enough to evaluate (4) for the case where $\mathbf{s}_1, \cdots, \mathbf{s}_4$, and $\mathbf{p}$ lie on the same line. The calculations are trivial and are omitted for brevity. □

It is worthwhile to note that any geometry in which the sensors and the target are almost collinear will also be problematic.

After establishing that there is a finite number of solutions for generic positions of the sensors where four doppler shift measurements are available, we present an answer to the second question posed in Problem 1. Before we formally propose the answer, note that the number of unknowns is four, and that with four pieces of data you get four polynomials equations, which have multiple solutions (though they may not all be real). However, with five pieces of data, one expects that the associated equations to have a unique solution. This solution is the one solution common to two selections of four. In the next proposition we prove that the position and the velocity of the target can be uniquely calculated if there are five doppler measurement available.

Proposition 3. *For $n \geq 5$ doppler measurements as described by* (2) *and generic positions for the sensors there is a unique solution for the unknown $\mathbf{x}$.*

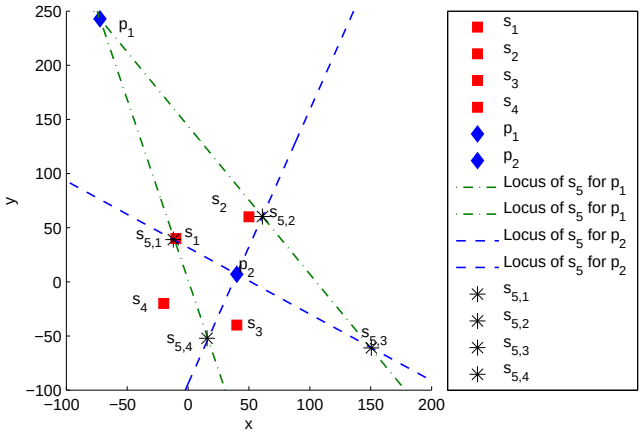


Fig. 1. Placing the fifth sensor at any of the positions indicated by * results in an ambiguous solution for the position of the target.

Proof: From Proposition 1 we know that for $n = 4$ there is a finite number of solutions for $\mathbf{x}$. Call the solutions for the position and the velocity of the target $\mathbf{x}_1, \cdots, \mathbf{x}_m$. Now, temporarily regard the position of sensor 5, $\mathbf{s}_5 = [s_{5,1}\ s_{5,2}]$ as an unknown. Consider the relationship between each of these solutions, $\mathbf{x}_i = [p_{i,1}\ p_{i,2}\ v_{i,1}\ v_{i,2}]^\top$, and the position of sensor 5, $\mathbf{s}_5$:

$$f_5^2\left((p_{i,1}-s_{5,1})^2+(p_{i,2}-s_{5,2})^2\right)- \\ \left(v_{i,1}(p_{i,1}-s_{5,1})+v_{i,2}(p_{i,2}-s_{5,2})\right)^2=0 \tag{10}$$

This equation results in two straight lines intersecting at $\mathbf{p}_i$, and one at least of which the true $\mathbf{s}_5$ must lie, namely that associated with the correct target position. We claim that generically only one such equation can be satisfied by the true $\mathbf{s}_5$. To establish the claim, we argue by contradiction. Assume that $\mathbf{x}_j$ and $\mathbf{x}_k$ define two different loci on which $\mathbf{s}_5$ must lie. These loci at most intersect at four points. So for all positions of $\mathbf{s}_5$ except these four points, $\mathbf{x}_j$ and $\mathbf{x}_k$ cannot be simultaneously the solutions to the localization problem via doppler measurements. With similar arguments one can eliminate any multiplicity of solutions for generic values of doppler shift measurements for $n \geq 5$. □

The proof of Proposition 3 indicates that, given four sensors in generic positions there are isolated positions where placing a fifth sensor does not lead to a unique solution for $\mathbf{x}$. An example of this scenario is depicted in Fig. 1, where two possible solutions or the location of the target are shown. Moreover, four positions are identified such that the placement of a fifth sensor at any of these positions will not resolve the ambiguity in the target position. An interesting direction for future work would be to consider the notion of optimal sensor placement as discussed in, e.g., [11], [12] which not only considers ambiguities but also the likely estimator performance given the relative sensor-target geometry.

So far, we have considered the case where the carrier frequency, f_c, is known. In some cases this is not realistic; in what comes next we show that there is no way from a single set of instantaneous doppler only measurements one can separate f_c and $\mathbf{v}$. We conclude this section by formalizing this fact and presenting a result on the case where the carrier frequency is unknown.

Proposition 4. *For $n \geq 5$ doppler measurements as described by* (2) *and generic positions for the sensors, and unknown carrier frequency f_c, there is a unique solution for the unknown $\mathbf{p}$ and for the vector $f_c\mathbf{v}$.*

Proof: Define $\nu = [\nu_1\ \nu_2]^\top \triangleq f_c\mathbf{v}$. For (2) we have

$$f_i = 2\frac{\nu^\top(\mathbf{p}-\mathbf{s}_i)}{c\|\mathbf{p}-\mathbf{s}_i\|} \tag{11}$$

From Proposition 3 we know that using five equations of the form (11), $\mathbf{p}$ and ν can be calculated uniquely. It follows easily that if one does not know f_c nor $\mathbf{v}$ and one estimates $\nu \triangleq f_c\mathbf{v}$ that for any chosen f_c (or $\mathbf{v}$) there is a subsequent value of $\mathbf{v}$ (or f_c) that satisfies $\nu = f_c\mathbf{v}$. Thus, f_c and $\mathbf{v}$ cannot be calculated separately.

Remark 1. *By using the calculated value $\mathbf{p}$ in consecutive time steps, one can estimate the velocity of the agent. Using this estimated velocity and knowing ν, one can further estimate f_c.*

III. Required Minimum Number of Hybrid Measurements

In this section we study the effect of having other types of measurements additional to doppler shift measurements on calculating the velocity and the position of the target.

Before continuing further, we note that to have a unique solution for a set of polynomial equations there is usually a need to have more equations than unknowns., save in cases where the equations are linear. However, having more equations than unknowns does not of itself guarantee the existence of a unique solution; the extra equation must in some way be independent, and it may have this property in almost all circumstances, i.e. generically, but not always. In this section we establish the cases where it can be mathematically shown that a unique solution for $\mathbf{x}$ exists when a combination of doppler and other measurements is available.

First, we consider the case where in addition to the doppler shift measurements described earlier, the distances between each of the sensors and the target can be measured as well. Denote the distance between the sensor i and the target as d_i where

$$d_i = \|\mathbf{p}-\mathbf{s}_i\|, \quad i = 1, \cdots, m \tag{12}$$

We have the following results:

Proposition 5. *For $n \geq 2$ doppler measurements as described by* (2) *and $m \geq 3$ distance measurements* (12) *and generic positions of the sensors there is a unique solution for the unknown $\mathbf{x}$.*

Proof: With three or more distance measurements the position of the target can be determined uniquely; i.e. three generic circles have at most a single point of intersection.

Knowing the position, then the doppler equations are linear equations in the velocity of the target. Hence, a unique solution for the velocity follows. □

Proposition 6. *For $n \geq 3$ doppler measurements as described by* (2) *and two distance measurements* (12) *and generic positions of the sensors there is a unique solution for the unknown* $\mathbf{x}$.

Proof: The two distance measurements pin down the target to a binary ambiguity; i.e. two generic circles have two points of intersection. Using each of these positions, then from the doppler equations we obtain two sets of three linear equations in the velocity of the target. Generically, only one of these sets forms a consistent system of linear equations. □

We now consider the case that instead of distance measurements, bearing measurements to the target at each of the sensors are available, viz.

$$\rho_i = [\rho_{i,1} \;\; \rho_{i,2}]^\top = \frac{\mathbf{p} - \mathbf{s}_i}{\|\mathbf{p} - \mathbf{s}_i\|} \tag{13}$$

is known at each sensor i. We formally consider this scenario in the next proposition.

Proposition 7. *For $n \geq 2$ doppler measurements as described by* (2) *and bearing measurements* (13) *there is a unique solution for the unknown* $\mathbf{x}$.

Proof: With at least two bearing measurements the position of the target can be determined uniquely. Knowing the position, then the doppler equations are linear equations in the velocity of the target. Hence, a unique solution for the velocity follows. □

In what comes next we consider the case where only one of the sensors is equipped with the capability to measure the target bearing. Without loss of generality assume that only sensor 1 can collect a bearing measurements to the target in addition to the doppler shift measurement. The rest of the sensors can only measure the doppler shift. For this scenario we have the following result.

Proposition 8. *For $n \geq 4$ doppler measurements as described by* (2) *and only one measured bearing to the target there is a unique solution for the unknown* $\mathbf{x}$.

Proof: Without loss of generality assume that the bearing measurement is measured at $\mathbf{s}_1$. For the doppler measurement at 1 we have

$$\begin{aligned} f_1 &= 2\mathbf{v}^\top \rho_1 \frac{f_c}{c} \\ f_1 &= 2(v_1\rho_{1,1} + v_2\rho_{1,2})\frac{f_c}{c} \end{aligned} \tag{14}$$

that is a linear equation in $\mathbf{v}$. Moreover, we have

$$p_2 = s_{1,2} + \rho_{1,2}\frac{p_1 - s_{1,1}}{\rho_{1,1}} \tag{15}$$

Calculating v_2 in terms of v_1 from (14) and p_2 in terms of p_1 from (14) and replacing them in

$$f_i = 2\mathbf{v}^\top \frac{\mathbf{p} - \mathbf{s}_i}{\|\mathbf{p} - \mathbf{s}_i\|}\frac{f_c}{c} \quad i = 2, 3, 4 \tag{16}$$

we obtain three quadratic equations in v_1 and p_1 only. Furthermore, it is known that generically three quadratic equations in two variables have a unique solution. Hence, $\mathbf{x}$ can be determined uniquely. □

In the next section, we introduce an algorithm to estimate the positon and the velocity of the target using the doppler-shift measurements measured at each of the sensor.

IV. An Algorithm to Estimate the Position and the Velocity of the Target

To calculate the position and the velocity of the target in the noiseless case it is enought to solve the following system of equations.

$$\delta_i - \frac{\mathbf{v}^\top(\mathbf{p} - \mathbf{s}_i)}{\|\mathbf{p} - \mathbf{s}_i\|} = 0 \quad i = 1, \ldots, n \tag{17}$$

where $\delta_i = \dfrac{cf_i}{2f_c}$. Equivalently due to the fact that the target and the sensors are not collocated (i.e. $||p - s_i||$ is nonzero), instead of solving (17) one can solve:

$$\delta_i\|\mathbf{p} - \mathbf{s}_i\| - \mathbf{v}^\top(\mathbf{p} - \mathbf{s}_i) = 0 \quad i = 1, \ldots, n, \tag{18}$$

or

$$\delta_i\sqrt{(\mathbf{p} - \mathbf{s}_i)^\top(\mathbf{p} - \mathbf{s}_i)} - \mathbf{v}^\top(\mathbf{p} - \mathbf{s}_i) = 0 \quad i = 1, \ldots, n. \tag{19}$$

Having the square-root in (19) makes it undesirable for solving numerically. Hence, instead we consider the following set of equations.

$$\delta_i^2(\mathbf{p} - \mathbf{s}_i)^\top(\mathbf{p} - \mathbf{s}_i) - (\mathbf{v}^\top(\mathbf{p} - \mathbf{s}_i))^2 = 0 \quad i = 1, \ldots, n. \tag{20}$$

Note that any solution of (19) is a solution of (20) but not vice versa. Assume $\mathbf{x}^\star = [\mathbf{p}^{\star\top} \;\; \mathbf{v}^{\star\top}]^\top$ is a solution to (19); then both $\mathbf{x}_1^\star = [\mathbf{p}^{\star\top} \;\; \mathbf{v}^{\star\top}]^\top$ and $\mathbf{x}_2^\star = [\mathbf{p}^{\star\top} \;\; -\mathbf{v}^{\star\top}]^\top$ satisfy (20). From Proposition 3 we know that for $n \geq 5$ (19) has a unique solution. Then it easily follows that for $n \geq 5$, (20) has exactly two solutions. Replacing $\mathbf{p}$ by $\mathbf{p}^\star$ in (17) results in a set of linear equations in $\mathbf{v}$ where only one of the values of $\mathbf{v}^\star$ or $-\mathbf{v}^\star$ satisfies it. The aforementioned analysis of the solutions for $\mathbf{x}$ in the noiseless case form the basis of the algorithm proposed in this section.

Now we consider the case where the doppler shift measurement is corrupted by noise. That is, each sensor measures $\hat{f}_i = f_i + w_i$ where w_i corresponds to the noise in the measurement carried out by sensor i. Setting $\hat{\delta}_i = \dfrac{c\hat{f}_i}{2f_c}$, we have

$$\hat{\delta}_i = \frac{\mathbf{v}^\top(\mathbf{p} - \mathbf{s}_i)}{\|\mathbf{p} - \mathbf{s}_i\|}. \tag{21}$$

Note that $\hat{\delta}_i = \delta_i + \omega_i$, where $\omega_i = \dfrac{cw_i}{2f_c}$. In the noisy case neither (19) nor (20) has a solution for $\mathbf{x}$. Instead we propose solving the following minimization problem to calculate the position and the velocity of the target. Solution of similar minimization problems when range, range-difference, and bearing measurments are available are studied in [13], [14].

$$[\mathbf{p}^\star, \mathbf{v}^\star] = \underset{\mathbf{p},\mathbf{v}}{\operatorname{argmin}} \;\; F(\mathbf{p}, \mathbf{v}), \tag{22}$$

where $F(\mathbf{p},\mathbf{v}) = \sum_{i=1}^{n}\left(\hat{\delta}_i^2\|\mathbf{p}-\mathbf{s}_i\|^2-(\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i))^2\right)^2$. The advantage of having such a cost function is that it is a polynomial in the unknowns, and can be minimized using modern polynomial optimization methods, e.g. see [15], [16]. By solving this minimization problem we obtain two values for the position and the velocity of the target, viz. $(\mathbf{p}_1^\star,\mathbf{p}_2^\star)$ and $(\mathbf{v}_1^\star,\mathbf{v}_2^\star)$, where $\mathbf{p}_1^\star=\mathbf{p}_2^\star$ and $\mathbf{v}_1^\star=-\mathbf{v}_2^\star$. To find the correct value for the velocity as before we replace $\mathbf{p}$ by $\mathbf{p}^\star\triangleq\mathbf{p}_1^\star=\mathbf{p}_2^\star$ in (21) to obtain a set of linear equations in $\mathbf{v}$:

$$\hat{\delta}_i=\frac{\mathbf{v}^\top(\mathbf{p}^\star-\mathbf{s}_i)}{\|\mathbf{p}^\star-\mathbf{s}_i\|}. \tag{23}$$

It is easy to check that the linear system of equations (23) is over-determined and inconsistent, and cannot be solved. However, a least-square solution to this syetem of linear equations can be obtained. The least squares solution to this system of equations is the estimated value for the target velocity, $\mathbf{v}^\star$. This procedure is outlined in Algorithm 1.

Algorithm 1 Target Position and Velocity Estimation Using n Doppler-shift Measurements Collecting at Nodes $(1,\ldots,n)$ in a Sensor Network Via Polynomial Optimization.
Input: $\hat{\delta}_1,\ldots,\hat{\delta}_n$
Output: $\mathbf{p}^\star$, $\mathbf{v}^\star$

Require: At least 5 sensors at generic positions.
$F(\mathbf{p},\mathbf{v})\leftarrow\sum_{i=1}^{n}\left(\hat{\delta}_i^2\|\mathbf{p}-\mathbf{s}_i\|^2-(\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i))^2\right)^2$
$[\mathbf{p}^\star,\ \mathbf{v}^\star]\leftarrow\underset{\mathbf{p},\mathbf{v}}{\text{argmin}}\ \ F(\mathbf{p},\mathbf{v})$
for $i=1:n$ **do**
$\hat{\rho}_i\leftarrow\dfrac{\mathbf{p}^\star-\mathbf{s}_i}{\|\mathbf{p}^\star-\mathbf{s}_i\|}$
end for
$\hat{A}\leftarrow[\hat{\rho}_1,\ldots,\hat{\rho}_1]^\top$
$\hat{\mathbf{b}}\leftarrow[\hat{\delta}_1,\ldots,\hat{\delta}_n]^\top$
$\mathbf{v}^\star\leftarrow\hat{A}^\dagger\hat{\mathbf{b}}$ $\{\hat{A}^\dagger$ is the pseudo-inverse of $\hat{A}.\}$
return $\mathbf{p}^\star$ and $\mathbf{v}^\star$

V. Illustrative Example

In this section we demonstrate the performance of the algorithm introduced in Section IV to estimate the position and the velocity of the target. We consider the setting depicted in Fig, 2. We consider the case where the measurements are corrupted by ten different levels of noise, where the noise is assumed to be gaussian with zero mean and variable variance (and independent at each sensor). The error in the estimates of position and the velocity for different noise levels after repeating the scenario for twenty times when six and seven measurements are used are depicted at Fig. 3 and Fig. 4 respectively. Moreover, note that the setting presented in Fig. 2 shows elements of a bad geometry as well: sensors at positions $\mathbf{s}_1$, $\mathbf{s}_2$, and $\mathbf{s}_3$ and the target are nearly collinear. However, here due to the presence of other sensors at non-collinear positions the estimation can be carried out effectievely.

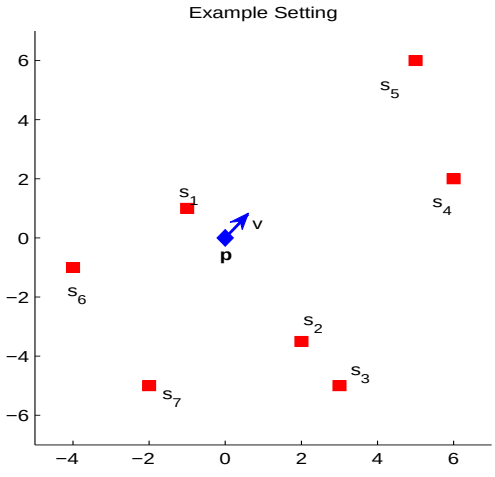


Fig. 2. The setting considered in the illustrative example. The squares denote the position of the sensors and the diamond is the target.

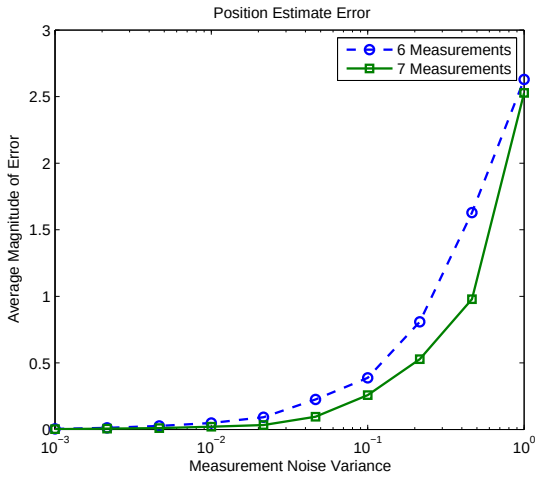


Fig. 3. The error in the estimate of the position of the target after repeating the scenario 20 times with variable noise variances when six and seven sensors are used.

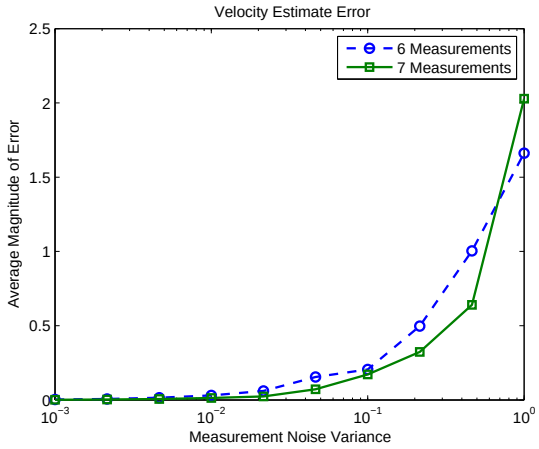


Fig. 4. The error in the estimate of the velocity of the target after repeating the scenario 20 times with variable noise variances when six and seven sensros are used.

VI. Concluding Remarks and Future Directions

The minimum number of doppler shift measurements necessary to have a finite number of solutions for the unknown target position and velocity is calculated analytically via algebraic arguments. Additionally, we stated the necessary and sufficient number of generic measurements to have a unique solution for the target parameters. Later, the same problem has been studied where in addition to doppler shift measurements, other types of measurements are available, e.g. bearing or distance to the target from each of the sensors. Finally, a method based on polynomial optimization is introduced to calculate an estimate for the position and the velocity of the target using noisy doppler shift measurements. A numerical example is presented to demonstrate the performance of this algorithm.

A possible future research direction is to design a dynamical estimator to estimate the position and the velocity of the target measuring the doppler shift measurements continuously. Alternatively, one could consider the notion of constraint-based optimization for localization as discussed in, e.g., [17]–[22].

References

[1] L.R. Malling. Radio doppler effect for aircraft speed measurements. *Proceedings of the IRE*, 35(11):1357–1360, November 1947.

[2] E.J. Barlow. Doppler radar. *Proceedings of the IRE*, 37(4):340–355, April 1949.

[3] Walter R. Fried. Principles and performance analysis of doppler navigation systems. *IRE Transactions on Aeronautical and Navigational Electronics*, 4(4):176–196, December 1957.

[4] S.N. Salinger and J.J. Brandstatter. Application of recursive estimation and kalman filtering to doppler tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 19(4):585–592, July 1970.

[5] D.C. Torney. Localization and Observability of Aircraft via Doppler Shifts. *Aerospace and Electronic Systems, IEEE Transactions on*, 43(3):1163–1168, 2007.

[6] Y.T. Chan and F.L. Jardine. Target localization and tracking from Doppler-shift measurements. *Oceanic Engineering, IEEE Journal of*, 15(3):251–257, 1990.

[7] M. Amin, P. Zemany, P. Setlur, and F. Ahmad. Moving target localization for indoor imaging using dual frequency CW radars. In *Sensor Array and Multichannel Processing, 2006. Fourth IEEE Workshop on*, pages 367–371. IEEE, 2008.

[8] Y.C. Xiao, P. Wei, and T. Yuan. Observability and Performance Analysis of Bi/Multi-Static Doppler-Only Radar. *Aerospace and Electronic Systems, IEEE Transactions on*, 46(4):1654–1667, 2010.

[9] N. Alam, A.T. Balaie, and A.G. Dempster. Dynamic path loss exponent estimation in a vehicular network using doppler effect and received signal strength. In *Proceedings of IEEE 71st Vehicular Technology Conference*, pages 1–5, 2010.

[10] B. Kusý, I. Amundson, J. Sallai, P. Völgyesi, A. Lédeczi, and X. Koutsoukos. Rf doppler shift-based mobile sensor tracking and navigation. *ACM Trans. Sen. Netw.*, 7:1:1–1:32, August 2010.

[11] A.N. Bishop, B. Fidan, B.D.O. Anderson, K. Dogancay, and P.N. Pathirana. Optimality analysis of sensor-target localization geometries. *Automatica*, 46(3):479–492, March 2010.

[12] A.N. Bishop and M. Smith. Remarks on the Cramer-Rao inequality for Doppler-based target parameter estimation. In *Proceedings of the Sixth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP'10)*, pages 199–204, Brisbane, Australia, December 2010.

[13] I. Shames, B. FIdan, B. D. O. Anderson, and H. Hmam. Cooperative self-localization of mobile agents. *To appear in IEEE Transactions on Aerospace and Electronic Systems*, 2011.

[14] I. Shames, P. T. Bibalan, B. Fidan, and B. D. O. Anderson. Polynomial methods in noisy network localization. In *17th Mediterranean Conference on Control and Automation*, pages 1307–1312. IEEE, 2009.

[15] D. Henrion and J. Lasserre. Detecting global optimality and extracting solutions in gloptipoly. In D. Henrion and A. Garulli, editors, *In Positive Polynomials in Control*, Lecture Notes on Control and Information Sciences. Springer-Verlag, 2005.

[16] P. Parrilo. Semidefinite programming relaxations for semi-algebraic problems. *Mathematical Programming*, 96(2):293–320, 2003.

[17] R.I. Hartley and P. Sturm. Triangulation. *Computer Vision and Image Understanding*, 68(2):146–157, November 1997.

[18] M. Cao, B.D.O. Anderson, and A.S. Morse. Localization with imprecise distance information in sensor networks. *Systems and Control Letters*, 55(11):887–893, November 2006.

[19] A.N. Bishop, B. Fidan, K. Dogancay, B.D.O. Anderson, and P.N. Pathirana. Exploiting geometry for improved hybrid AOA/TDOA based localization. *Signal Processing*, 88(7):1775–1791, July 2008.

[20] A.N. Bishop, B. Fidan, B.D.O. Anderson, K. Dogancay, and P.N. Pathirana. Optimal range-difference-based localization considering geometrical constraints. *IEEE Journal of Oceanic Engineering*, 33(3):289–301, July 2009.

[21] A.N. Bishop, B. Fidan, B.D.O. Anderson, P.N. Pathirana, and G. Mao. Bearing-only localization using geometrically constrained localization. *IEEE Transactions on Aerospace and Electronic Systems*, 45(1):308–320, January 2009.

[22] A.N. Bishop. A tutorial on constraints for positioning on the plane. In *Proceedings of the 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC'10)*, pages 1689–1694, Istanbul, Turkey, 2010.

Doppler Shift Target Localization

Iman Shames, *Member, IEEE,* Adrian N. Bishop, *Member, IEEE,* Matthew Smith, *Member, IEEE,* Brian D. O. Anderson, *Life Fellow, IEEE,*

Abstract

This paper outlines the problem of doppler-based target position and velocity estimation using a sensor network. The minimum number of doppler shift measurements at distinct generic sensor positions to have a finite number of solutions, and later, a unique solution for the unknown target position and velocity is stated analytically. Furthermore, we study the same problem where not only doppler shift measurements are collected, but also other types of measurements are available, e.g. bearing or distance to the target from each of the sensors. Later, we study the Cramer-Rao inequality associated with the doppler-shift measurements to a target in a sensor network, and use the Cramer-Rao bound to illustrate some results on optimal placements of the sensors when the goal is to estimate the velocity of the target. Some simulation results are presented in the end.

Index Terms

Doppler Measurements, Location Estimation, Doppler Localization, Cramer-Rao Inequality, Fisher Information Matrix

I. Introduction

Using doppler-shifts for position and velocity estimation has a long history; see e.g. [1]–[8]. Recently, the doppler effect has gained a renewed interest and it has been implemented for cooperative positioning in vehicular networks [9].

In this paper, we consider a scenario with n nodes with both transmitting and sensing capabilities, that are called sensors for the rest of this paper. The target has an unknown position and velocity $\mathbf{x} = [\mathbf{p}^\top \ \mathbf{v}^\top]^\top \in \mathbb{R}^4$. The position of the non-collocated sensors is given by $\mathbf{s}_i = [s_{i,1} \ s_{i,2}]^\top \in \mathbb{R}^2$, $\forall i \in \{1, \dots, n\}$.

The measured doppler-shift is $\hat{\delta}_i$ at the i^{th} sensor and is caused by a target reflection due to a signal generated earlier by the same sensor. This frequency shift can be approximated by [10]

$$\hat{\delta}_i = \delta_i + w_i \tag{1a}$$

$$= 2\frac{f_{c,i}}{c}\left(\frac{(\mathbf{p}-\mathbf{s}_i)^\top}{\|\mathbf{p}-\mathbf{s}_i\|}\right)\mathbf{v} + w_i \tag{1b}$$

where c is the speed of light (or signal propagation) and $\|\cdot\|$ is the standard Euclidean vector norm and $f_{c,i}$ is the carrier frequency employed by this sensor. Finally, w_i is the noise variable. Note here that the localization is to be achieved instantaneously;

I. Shames is with the ACCESS Linnaeus Centre, Royal Institute of Technology (KTH) in Stockholm, Sweden, e-mail: imansh@kth.se. A.N. Bishop and B.D.O. Anderson are with NICTA, Canberra Research Lab and the Australian National University (ANU), e-mail: {Adrian.Bishop,Brian.Anderson}@anu.edu.au. M. Smith is with CEA Technologies, Canberra, Australia, e-mail: Matthew.Smith@cea.com.au. I. Shames is supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation. A.N. Bishop and B.D.O. Anderson are supported by USAF-AOARD-10-4102 and by NICTA which is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. A.N. Bishop is also supported by the Australian Research Council (ARC) via a Discovery Early Career Researcher Award (DE-120102873). M. Smith is supported by CEA Technologies.

we are not envisaging collecting information from agents at a number of successive instants of time and using them to infer position at a single instant of time (The connection with filtering methods is explored further below). There are many studies in the literature that try to solve a similar problem via collecting measurements over a time interval and feeding them into an estimator. For example, in [5] the problem of localization of a single aircraft using doppler measurements is studied. Other similar approaches can be found in [4], [6]–[8]. The analysis carried out in this paper along with the optimization method proposed can be considered as constituting a batch processing method for instantaneous estimation of the target location and velocity. This in turn might be used to initialize and improve the updates of any other implemented filter which tracks the target position as the target moves in the environment. For example, Kalman-based filters are prone to errors when the target's motion model deviates significantly from the actual target motion. Our analysis can guard against such behavior and re-initialize the filter, e.g. see [11] and references therein.

The first main contribution of this paper is that the minimum number of doppler shift measurements required to have a finite number of solutions for the unknown $\mathbf{x}$ is algebraically derived. In some scenarios, a separate piece of knowledge about the target will allow disambiguation. Later, the result is extended to the case where having a unique solution is required. Moreover, the scenarios where different types of measurements, e.g. direction-of-arrival, or distance, are available in addition to doppler shift measurements are considered. The aforementioned conclusions assume zero measurement noise; following on from this, an optimization method based on polynomial optimization methods is introduced to calculate the velocity and the position of a target where noisy doppler shift measurements are available. Later, we calculate the Cramer-Rao inequality for different sensor network configurations and present some discussions of the Fisher Information matrix, which in turn leads to the introduction of optimal sensor placement in the scenarios where a network of sensors gather doppler-shift measurements in the signal from a target.

The remaining sections of this paper are organized as follows. In the next section the main problem of interest in considered. In Section III the case where different types of measurement in addition to doppler shift measurement are available to the sensors is considered. A method based on polynomial optimization to estimate the position and the velocity of the target where doppler shift measurements are contaminated by noise is presented in Section IV. The Cramer-Rao inequality for the scenario considered here is calculated in Section V. In Section VI some insights into the Fisher Information matrix for the velocity estimation are presented. An illustrative example presenting the performance of the proposed optimization method is given in Section VII. Concluding remarks and future directions come in Section VIII.

II. Required Minimum Number of Doppler Shift Measurements

Initially, it is assumed that the measurements are noiseless, i.e. $w_i = 0$.

$$\begin{aligned}\hat{\delta}_i &= \delta_i \\ &= 2\frac{\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i)}{\|\mathbf{p}-\mathbf{s}_i\|}\frac{f_{c,i}}{c}.\end{aligned} \tag{2}$$

Normalizing so that $f_i \triangleq \frac{c\delta_i}{2f_{c,i}}$, we obtain

$$f_i = \frac{\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i)}{\|\mathbf{p}-\mathbf{s}_i\|} \tag{3}$$

Now we are ready to pose the problem of interest in this section.

Problem 1. *Consider n stationary sensors at $\mathbf{s}_i \in \mathbb{R}^2$ capable of collecting noiseless doppler shift measurements from a target at position $\mathbf{p}$ moving with a nonzero velocity $\mathbf{v}$ of the form* (2)*.*

1) *What is the minimum value for n such that there is a finite number of solutions for $\mathbf{x}$?*
2) *What is the minimum value for n such that there is a unique solution for $\mathbf{x}$?*

We limit our analysis to the case where the nodes and the target are in $\mathbb{R}^2$, however, note that the analysis for the case that they are in $\mathbb{R}^3$ is much the same. First we have the following remark.

Remark 1. *Throughout this paper, when it is stated that a property is held for generic positions of the sensors, it is meant that such a property holds for all positions of the sensors except for sensor positions in a set of measure zero.*

The answer to the first question posed in Problem 1 is formally presented in the following proposition for the case where $n = 4$ doppler shift measurements are available. This proposition states that with $n = 4$ measurements generically [1] we have a finite number of solutions and one might be able to disambiguate the solutions using other measurements, e.g. range, or bearing measurements. Moreover, disambiguation may also be made possible due to prior measurements, or to a priori knowledge about the geographic constraints on targets in the area of interest. This particularly is important for the cases where the possible solutions are widely separated, e.g., see Fig. 1. Hence this situation is of potential practical interest.

Proposition 1. *For $n = 4$ doppler measurements as described by* (2) *and generic positions of the sensors there is a finite number of solutions for the unknown $\mathbf{x}$.*

Proof: Denote the noiseless mapping from the agent position and velocity, $\mathbf{x} = [\mathbf{p}^\top\ \mathbf{v}^\top]^\top$ (a vector in $\mathbb{R}^4$) to measurements (another vector in $\mathbb{R}^n$) by F, where n is the number of sensors. More specifically $F(\mathbf{x}) = [f_1\ f_2\ f_3\ f_4]^\top$. Let ∇F denote the Jacobian of F:

$$\nabla F = \begin{bmatrix} \frac{\partial f_1}{\partial p_1} & \frac{\partial f_1}{\partial p_2} & \frac{\partial f_1}{\partial v_1} & \frac{\partial f_1}{\partial v_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_4}{\partial p_1} & \frac{\partial f_4}{\partial p_2} & \frac{\partial f_4}{\partial v_1} & \frac{\partial f_4}{\partial v_2} \end{bmatrix} \tag{4}$$

where

$$\frac{\partial f_i}{\partial p_1} = \frac{-(p_2 - s_{i,2})(s_{i,2}v_1 - s_{i,1}v_2 + p_1v_2 - p_2v_1)}{\|\mathbf{s}_i - \mathbf{p}\|^3} \tag{5}$$

$$\frac{\partial f_i}{\partial p_2} = \frac{(p_1 - s_{i,1})(s_{i,2}v_1 - s_{i,1}v_2 + p_1v_2 - p_2v_1)}{\|\mathbf{s}_i - \mathbf{p}\|^3} \tag{6}$$

[1]There will be special positions of sensors for which $\mathbf{x}$ cannot be determined–for example if they are all collocated. However, the set of such exceptional positions is a set of measure zero; the word 'generically' captures this notion.

$$\frac{\partial f_i}{\partial v_1} = \frac{(p_1 - s_{i,1})}{\|\mathbf{s}_i - \mathbf{p}\|} \tag{7}$$

$$\frac{\partial f_i}{\partial v_2} = \frac{(p_2 - s_{i,2})}{\|\mathbf{s}_i - \mathbf{p}\|} \tag{8}$$

It is easy to check that for generic values of $\mathbf{x}$, $\mathbf{s}_i$, $i = 1, \cdots, 4$, ∇F is not singular. Moreover, we know that the set of solutions to (2) form an algebraic variety , [12], i.e. the solution set can be defined by solving a set of multivariable polynomial equations. The reason for this is that while (2) is not a polynomial equation, it is easy to see that its zeros are also the zeros of the following polynomial equation.

$$f_i^2 \|\mathbf{p} - \mathbf{s}_i\|^2 - (\mathbf{v}^\top (\mathbf{p} - \mathbf{s}_i))^2 = 0 \tag{9}$$

The set of the solutions to the equations described by (9) is known to have at least one member; that is the solution corresponding to the physical setup. The nonsingularity of the Jacobian implies that for generic values for measurements we do not have a continuous set of solutions. As a result, the variety is a zero-dmensional variety [12]. Further, any zero-dimensional variety has a finite number of points [12], and so the solution set as a subset of a zero-dimensional variety also has a finite number of points, i.e., there is a finite number of solutions for the localization problem using doppler measurements where $n \geq 4$. ■

Now we briefly consider the case where the Jacobian matrix ∇F is singular. This corresponds to those sensor and target geometries where there is an infinite number of solutions for the unknown $\mathbf{p}$ and $\mathbf{v}$. We call these geometries as *degenerate geometries*. The following proposition characterizes two of these degenerate geometries.

Proposition 2. *The Jacobian matrix ∇F given in* (4) *is singular in the following situations:*

1) *If any pair of the sensors $\mathbf{s}_1, \cdots, \mathbf{s}_4$ and the target $\mathbf{p}$ are collinear.*
2) *If $\mathbf{v} = 0$.*

Proof: To prove the first statement of this proposition it is enough to evaluate (4) for the case where two of the sensors $\mathbf{s}_1, \cdots, \mathbf{s}_4$, and $\mathbf{p}$ lie on the same line. The calculations are trivial and are omitted for brevity. For establishing the validity of the second statement it is enough to observe that the first two columns of (4) are zero when $\mathbf{v} = 0$. ■

It is worthwhile to note that in practice any geometry in which the sensors and the target are almost collinear will also be problematic due to the presence of noise in the measurements.

After establishing that there is a finite number of solutions for generic positions of the sensors where four doppler shift measurements are available, we present an answer to the second question posed in Problem 1. Before we formally propose the answer, note that the number of unknowns is four, and that with four pieces of data we have four polynomials equations, which in general have a finite set multiple solutions (though they may not all be real). However, with five pieces of data, one expects that the associated equations generically have a unique solution. This solution is the one solution common to two selections of four. In the next proposition we prove that the position and the velocity of the target can be uniquely calculated if there are five doppler measurement available.

Proposition 3. *For $n \geq 5$ doppler measurements as described by* (2) *and generic positions for the sensors there is a unique*

Insert Figure 1 here

Fig. 1. Placing the fifth sensor at any of the positions indicated by * results in an ambiguous solution for the position of the target.

solution for the unknown $\mathbf{x}$.

Proof: From Proposition 1 we know that for $n = 4$ there is a finite number of solutions for $\mathbf{x}$, m say. Call the solutions for the position and the velocity of the target $\mathbf{x}_1, \cdots, \mathbf{x}_m$. Now, temporarily regard the position of sensor 5, $\mathbf{s}_5 = [s_{5,1}\ s_{5,2}]$ as an unknown. Consider the relationship between each of these solutions, $\mathbf{x}_i = [p_{i,1}\ p_{i,2}\ v_{i,1}\ v_{i,2}]^\top$, and the position of sensor 5, $\mathbf{s}_5$:

$$\begin{aligned} f_5^2\left((p_{i,1}-s_{5,1})^2+(p_{i,2}-s_{5,2})^2\right)- \\ \left(v_{i,1}(p_{i,1}-s_{5,1})+v_{i,2}(p_{i,2}-s_{5,2})\right)^2=0 \end{aligned} \tag{10}$$

This equation (regarded as an equation for the two temporarily unknown coordinates $s_{5,1}$ and $s_{5,2}$, results in two straight lines intersecting at $\mathbf{p}_i$; the true $\mathbf{s}_5$ must lie on one of these straight line pairs, namely that associated with the correct target position. We claim that generically only one such equation can be satisfied by the true $\mathbf{s}_5$. To establish the claim, we argue by contradiction. Assume that $\mathbf{x}_j$ and $\mathbf{x}_k$ define two different loci on both of which $\mathbf{s}_5$ lies. These loci at most intersect at four points. So for all positions of $\mathbf{s}_5$ except these four points, $\mathbf{x}_j$ and $\mathbf{x}_k$ cannot be simultaneously the solutions to the localization problem via doppler measurements. With similar arguments one can eliminate any multiplicity of solutions for generic values of doppler shift measurements for $n \geq 5$. ■

The proof of Proposition 3 indicates that, given four sensors in generic positions there are isolated positions where placing a fifth sensor does not lead to a unique solution for $\mathbf{x}$. An example of this scenario is depicted in Fig. 1, where two possible straight line pairs are shown, being determined by two of the possible finite target positions computed using the measurements from sensors 1 to 4. Moreover, four positions are identified such that the placement of a fifth sensor at any of these positions will not resolve the ambiguity in the target position.

So far, we have considered the case where each sensor i emits a signal with a known carrier frequency $f_{c,i}$. In what comes next we consider two other cases. First, we consider a scenario that occurs commonly. In this scenario an illuminating emitter and a receiver are colocated and all the other sensors only receive the reflection of the signal emitted by the first sensor off the target. Suppose the illuminating radar and the first receiver are colocated at $\mathbf{s}_1$ and the remaining sensors are located at $\mathbf{s}_2, \ldots, \mathbf{s}_\ell$. Let f_c be the frequency of the illuminating radar. Then for $i = 1$ there holds

$$\delta_1 = 2\frac{f_c}{c}\frac{\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_1)}{\|\mathbf{p}-\mathbf{s}_1\|}, \tag{11}$$

and for $i = 2, \ldots, \ell$, [10]

$$\delta_i = \frac{f_c}{c}\left[\frac{\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_1)}{\|\mathbf{p}-\mathbf{s}_1\|}+\frac{\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i)}{\|\mathbf{p}-\mathbf{s}_i\|}\right]. \tag{12}$$

We note that from this data, it is trivial to compute the set of values

$$\bar{f}_i = \frac{\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i)}{\|\mathbf{p}-\mathbf{s}_i\|}$$

for $i = 1, 2, \ldots, \ell$, where

$$\bar{f}_i \triangleq \frac{c(2\delta_i - \delta_1)}{2f_c} \quad i = 1, 2, \ldots, \ell.$$

In other words, from a purely mathematical point of view, this scenario can be reduced to the original one, where all the $f_{c,i}$ assume the same value. Hence, the same analysis applies to this scenario.

Second, we consider the case where the target emits a signal with the frequency f_c and the sensors just measure this frequency without emitting any signal of their own. This scenario is usually known as a passive doppler localization scenario. In this case the noiseless measurements are of the form

$$\delta_i = \frac{f_c}{c}\frac{\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i)}{\|\mathbf{p}-\mathbf{s}_i\|} \tag{13}$$

The theory applying when f_c is independently known a priori is effectively the same as the one stated previously. However, if f_c is unknown, the question arises as to whether it can be determined. For this scenario, we show that there is no way from a single set of instantaneous doppler only measurements that one can separate f_c and $\mathbf{v}$. We conclude this section by formalizing this fact and presenting a result on the case where the carrier frequency is unknown.

Proposition 4. *For $n \geq 5$ doppler measurements as described by* (13) *and generic positions for the sensors, and unknown carrier frequency f_c, there is a unique solution for the unknown $\mathbf{p}$ and for the vector $f_c\mathbf{v}$.*

Proof: Define $\nu = [\nu_1\ \nu_2]^\top \triangleq f_c\mathbf{v}$. For (13) we have

$$\delta_i = \frac{\nu^\top(\mathbf{p}-\mathbf{s}_i)}{c\|\mathbf{p}-\mathbf{s}_i\|} \tag{14}$$

From Proposition 3 we know that using five equations of the form (14), $\mathbf{p}$ and ν can be calculated uniquely. It follows easily that if one does neither know f_c nor $\mathbf{v}$ and only estimates $\nu \triangleq f_c\mathbf{v}$ then for any chosen f_c (or $\mathbf{v}$) there is a subsequent value of $\mathbf{v}$ (or f_c) that satisfies $\nu = f_c\mathbf{v}$. Thus, f_c and $\mathbf{v}$ cannot be calculated separately. ■

Remark 2. *By using the calculated value $\mathbf{p}$ in consecutive time steps, one can estimate the velocity of the agent. Using this estimated velocity and knowing ν, one can further estimate f_c.*

III. Required Minimum Number of Hybrid Measurements

In this section we study the effect of having other types of measurements additional to doppler shift measurements on calculating the velocity and the position of the target.

Before continuing further, we note that to have a *unique* solution for a set of polynomial equations there is usually a need to have *more* equations than unknowns, save in cases where the equations are linear. However, having more equations than

unknowns does not of itself guarantees the existence of a unique solution; the extra equation must in some way be independent, and it may have this property in almost all circumstances, i.e. generically, even if not always. In this section we establish the cases where it can be mathematically shown that a unique solution for $\mathbf{x}$ exists when a combination of doppler and other measurements is available. We have the following results:

Proposition 5. *The following statements are true.*

1) *First, consider the case where in addition to the doppler shift measurements described earlier, the distances between each of the sensors and the target can be measured as well. Denote the distance between the sensor i and the target as d_i where*

$$d_i = \|\mathbf{p} - \mathbf{s}_i\|, \quad i = 1, \cdots, m_d \tag{15}$$

   a) *For $n \geq 2$ doppler measurements as described by* (2) *and $m_d \geq 3$ distance measurements* (15) *and generic positions of the sensors there is a unique solution for the unknown $\mathbf{x}$.*
   b) *For $n \geq 3$ doppler measurements as described by* (2) *and $m_d \geq 2$ distance measurements* (15) *and generic positions of the sensors there is a unique solution for the unknown $\mathbf{x}$.*

2) *Second, consider the case that instead of distance measurements, bearing measurements to the target at each sensor i are available, viz.*

$$\psi_i = [\psi_{i,1} \;\; \psi_{i,2}]^\top = \frac{\mathbf{p} - \mathbf{s}_i}{\|\mathbf{p} - \mathbf{s}_i\|}, \quad i = 1, \cdots, m_b \tag{16}$$

   is known at each sensor i.

   a) *For $n \geq 2$ doppler measurements as described by* (2) *and at $m_b \geq 2$ bearing measurements* (16) *and generic positions of the sensors there is a unique solution for the unknown $\mathbf{x}$.*
   b) *For $n \geq 4$ doppler measurements as described by* (2) *and only $m_b = 1$ measured bearing to the target and generic positions of the sensors there is a unique solution for the unknown $\mathbf{x}$.*

3) *Third, for $n \geq 3$ doppler measurements as described by* (2)*, at least $m_d = 1$ distance measurement* (15)*, at least $m_b = 1$ measured bearing to the target* (16)*, and generic positions of the sensors there is a unique solution for the unknown $\mathbf{x}$.*

Proof: The proof of 1a is trivial.

To prove 1b observe that the two distance measurements pin down the target to a binary ambiguity; i.e. two generic circles have two points of intersection. Using each of these positions, then from the doppler equations we obtain two sets of three linear equations in the velocity of the target. Generically, only one of these sets forms a consistent system of linear equations.

The proof of 2a is obvious.

To prove 2b without loss of generality assume that only sensor 1 can collect a bearing measurements to the target in addition to the doppler shift measurement. The rest of the sensors can only measure the doppler shift. For the doppler measurement at

TABLE I
MINIMUM NUMBER OF MEASUREMENTS IN DIFFERENT SCENARIOS FOR UNIQUE TARGET POSITION AND VELOCITY LOCALIZATION

No. of Doppler Measurements	5	2	3	2	4	3
No. of Distance Measurements	0	3	2	0	0	1
No. of Bearing Measurements	0	0	0	2	1	1

1 we have

$$\begin{aligned} \delta_1 &= 2\mathbf{v}^\top \psi_1 \frac{f_{c,i}}{c} \\ \delta_1 &= 2(v_1\psi_{1,1} + v_2\psi_{1,2})\frac{f_{c,i}}{c} \end{aligned} \tag{17}$$

that is a linear equation in $\mathbf{v}$. Moreover, we have

$$p_2 = s_{1,2} + \psi_{1,2}\frac{p_1 - s_{1,1}}{\psi_{1,1}} \tag{18}$$

Calculating v_2 in terms of v_1 from (17) and p_2 in terms of p_1 from (18) and replacing them in

$$\delta_i = 2\mathbf{v}^\top \frac{\mathbf{p} - \mathbf{s}_i}{\|\mathbf{p} - \mathbf{s}_i\|}\frac{f_{c,i}}{c} \quad i = 2, 3, 4 \tag{19}$$

we obtain three quadratic equations in v_1 and p_1 only. Furthermore, it is known that generically three quadratic equations in two variables have a unique solution. Hence, $\mathbf{x}$ can be determined uniquely.

The proof of the last statement is very similar to that of 1b and is omitted. ■

Table I summarises the various cases we have considered. In the next section, we introduce an algorithm to estimate the position and the velocity of the target using the doppler-shift measurements measured at each of the sensor, where noise may contaminate the measurements.

IV. AN ALGORITHM TO ESTIMATE THE POSITION AND THE VELOCITY OF THE TARGET

To calculate the position and the velocity of the target in the noiseless case it is enough to solve the following system of equations.

$$f_i - \frac{\mathbf{v}^\top(\mathbf{p} - \mathbf{s}_i)}{\|\mathbf{p} - \mathbf{s}_i\|} = 0 \quad i = 1, \ldots, n \tag{20}$$

where $f_i = \dfrac{c\delta_i}{2f_{c,i}}$. Equivalently due to the fact that the target and the sensors are not collocated (i.e. $||\mathbf{p} - \mathbf{s}_i||$ is nonzero), instead of solving (20) one can solve:

$$f_i\|\mathbf{p} - \mathbf{s}_i\| - \mathbf{v}^\top(\mathbf{p} - \mathbf{s}_i) = 0 \quad i = 1, \ldots, n, \tag{21}$$

or

$$f_i\sqrt{(\mathbf{p} - \mathbf{s}_i)^\top(\mathbf{p} - \mathbf{s}_i)} - \mathbf{v}^\top(\mathbf{p} - \mathbf{s}_i) = 0 \quad i = 1, \ldots, n. \tag{22}$$

Having the square-root in (22) makes it undesirable for solving numerically. Hence, instead we consider the following set of equations.

$$f_i^2(\mathbf{p}-\mathbf{s}_i)^\top(\mathbf{p}-\mathbf{s}_i)-(\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i))^2=0 \quad i=1,\ldots,n. \tag{23}$$

Note that any solution of (22) is a solution of (23) but not vice versa. Assume $\mathbf{x}^\star=[\mathbf{p}^{\star\top} \;\; \mathbf{v}^{\star\top}]^\top$ is a solution to (22); then both $\mathbf{x}_1^\star=[\mathbf{p}^{\star\top} \;\; \mathbf{v}^{\star\top}]^\top$ and $\mathbf{x}_2^\star=[\mathbf{p}^{\star\top} \;\; -\mathbf{v}^{\star\top}]^\top$ satisfy (23). From Proposition 3 we know that for $n\geq 5$ (22) has a unique solution. Then it easily follows that for $n\geq 5$, (23) has exactly two solutions. Replacing $\mathbf{p}$ by $\mathbf{p}^\star$ in (20) results in a set of linear equations in $\mathbf{v}$ where only one of the values of $\mathbf{v}^\star$ or $-\mathbf{v}^\star$ satisfies it. The aforementioned analysis of the solutions for $\mathbf{x}$ in the noiseless case form the basis of the algorithm proposed in this section.

Now we consider the case where the doppler shift measurement is corrupted by noise. That is, each sensor measures $\hat{\delta}_i=\delta_i+w_i$ where w_i corresponds to the noise in the measurement carried out by sensor i. Setting $\hat{f}_i=\dfrac{c\hat{\delta}_i}{2f_{c,i}}$, we have

$$\hat{f}_i=\frac{\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i)}{\|\mathbf{p}-\mathbf{s}_i\|}. \tag{24}$$

Note that $\hat{f}_i=f_i+\omega_i$, where $\omega_i=\dfrac{cw_i}{2f_{c,i}}$. In the noisy case neither (22) nor (23) have a solution for $\mathbf{x}$. Instead we propose solving the following minimization problem to calculate the position and the velocity of the target. Solution of similar minimization problems when range, range-difference, and bearing measurments are available are studied in [13], [14].

$$[\mathbf{p}^\star,\mathbf{v}^\star]=\underset{\mathbf{p},\mathbf{v}}{\text{argmin}} \;\; J(\mathbf{p},\mathbf{v}), \tag{25}$$

where $J(\mathbf{p},\mathbf{v})=\sum\limits_{i=1}^{n}\left(\hat{f}_i^2\|\mathbf{p}-\mathbf{s}_i\|^2-(\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i))^2\right)^2$. The advantage of having such a cost function is that it is a polynomial in the unknowns, and can be minimized using modern polynomial optimization methods, e.g. see [15], [16]. By solving this minimization problem we obtain two values for the position and the velocity of the target, viz. $(\mathbf{p}_1^\star,\mathbf{p}_2^\star)$ and $(\mathbf{v}_1^\star,\mathbf{v}_2^\star)$, where $\mathbf{p}_1^\star=\mathbf{p}_2^\star$ and $\mathbf{v}_1^\star=-\mathbf{v}_2^\star$. To find the correct value for the velocity as before we replace $\mathbf{p}$ by $\mathbf{p}^\star\triangleq\mathbf{p}_1^\star=\mathbf{p}_2^\star$ in (24) to obtain a set of linear equations in $\mathbf{v}$:

$$\hat{f}_i=\frac{\mathbf{v}^\top(\mathbf{p}^\star-\mathbf{s}_i)}{\|\mathbf{p}^\star-\mathbf{s}_i\|}. \tag{26}$$

It is easy to check that the linear system of equations (26) is over-determined and inconsistent, and cannot be solved. However, a least-square solution to this syetm of linear equations can be obtained. The least squares solution to this system of equations is the estimated value for the target velocity, $\mathbf{v}^\star$. This procedure is outlined in Algorithm 1.

We conclude this section by briefly considering the maximum likelihood problem of

$$[\mathbf{p}_{ML}^\star,\mathbf{v}_{ML}^\star]=\underset{\mathbf{p},\mathbf{v}}{\text{argmin}} \;\; J_{ML}(\mathbf{p},\mathbf{v}), \tag{27}$$

where $J_{ML}(\mathbf{p},\mathbf{v})=\sum\limits_{i=1}^{n}\left(\hat{f}_i-\dfrac{\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i)}{\|\mathbf{p}-\mathbf{s}_i\|}\right)^2$. Solving (27) involves solving a nonlinear least squares problem that its accuracy depends on the initial guess for its solution. We note that, the output of Algorithm 1 can be used to initialize the nonlinear

Algorithm 1 Target Position and Velocity Estimation Using n doppler-shift Measurements Collecting at Nodes $(1,\ldots,n)$ in a Sensor Network Via Polynomial Optimization.
Input: $\hat{f}_1,\ldots,\hat{f}_n$
Output: $\mathbf{p}^\star$, $\mathbf{v}^\star$

Require: At least 5 sensors at generic positions.
$J(\mathbf{p},\mathbf{v}) \leftarrow \sum_{i=1}^{n}\left(\hat{f}_i^2\|\mathbf{p}-\mathbf{s}_i\|^2-(\mathbf{v}^\top(\mathbf{p}-\mathbf{s}_i))^2\right)^2$
$[\mathbf{p}^\star,\ \mathbf{v}^\star] \leftarrow \underset{\mathbf{p},\mathbf{v}}{\text{argmin}}\ \ J(\mathbf{p},\mathbf{v})$
for $i=1:n$ **do**
$\hat{\psi}_i \leftarrow \dfrac{\mathbf{p}^\star-\mathbf{s}_i}{\|\mathbf{p}^\star-\mathbf{s}_i\|}$
end for
$\hat{A} \leftarrow [\hat{\psi}_1,\ldots,\hat{\psi}_n]^\top$
$\hat{\mathbf{b}} \leftarrow [\hat{f}_1,\ldots,\hat{f}_n]^\top$
$\mathbf{v}^\star \leftarrow \hat{A}^\dagger\hat{\mathbf{b}}$ $\{\hat{A}^\dagger$ is the pseudo-inverse of $\hat{A}.\}$
return $\mathbf{p}^\star$ and $\mathbf{v}^\star$

solver that minimizes (27). We demonstrate this later in the paper.

After establishing the minimum number of doppler measurements necessary to achieve localization, in the following sections we study the Cramer-Rao inequality for calculating the target position and velocity. Later, we propose an optimal sensor placement where the objective is to estimate the velocity of the target.

V. The Cramer-Rao Inequality

If $\mathcal{I}(\mathbf{x})$ is the Fisher information matrix, that will be formally defined below, then the Cramer-Rao inequality lower bounds the variance achievable by an unbiased estimator. For an *unbiased* estimate $\widehat{\mathbf{x}}$ of $\mathbf{x}$ we find

$$\mathrm{E}\left[(\widehat{\mathbf{x}}-\mathbf{x})(\widehat{\mathbf{x}}-\mathbf{x})^\top\right] \geq \mathcal{I}(\mathbf{x})^{-1} \tag{28}$$

If $\mathcal{I}(\mathbf{x})$ is singular then (in general) no unbiased estimator for $\mathbf{x}$ exists with a finite variance [17]. If (28) holds with equality, for some unbiased estimate $\widehat{\mathbf{x}}$, then the estimator is called *efficient* and the parameter estimate $\widehat{\mathbf{x}}$ is unique [17]. However, even if $\mathcal{I}(\mathbf{x})$ is non-singular then it is not practically guaranteed that an unbiased estimator can be recognized. Alternatively, if an unbiased estimator can be realized, it is not guaranteed that an efficient estimator exists [18].

The condition (28) says nothing about the performance and realizability of biased estimators. That is, in order to use (28) we must consider only unbiased estimators [17]. The $(i,j)^{th}$ element of $\mathcal{I}(\mathbf{x})$ is given by [19]

$$\mathcal{I}_{i,j}(\mathbf{x}) = \mathrm{E}\left[\frac{\partial}{\partial x_i}\ln\left(f_{\widehat{\mathbf{p}}}(\widehat{\mathbf{p}};\mathbf{x})\right)\frac{\partial}{\partial x_j}\ln\left(f_{\widehat{\mathbf{p}}}(\widehat{\mathbf{p}};\mathbf{x})\right)\right] \tag{29}$$

where $[\mathbf{p}^\top\ \mathbf{v}^\top]^\top = [x_1\ \ldots\ x_4] \in \mathbb{R}^4$ and $f_{\mathbf{f}}(\mathbf{f};\mathbf{x})$ is the Gaussian likelihood function. We then easily find $\mathcal{I}(\mathbf{x}) = \nabla F^\top\mathbf{R}_{\mathbf{f}}^{-1}\nabla F$. The Fisher information metric characterizes the nature of the likelihood function. If the likelihood function is sharply peaked then the true parameter value is easier to estimate from the measurements than if the likelihood function is flatter.

A. General Results

The Fisher information matrix is given by (30).

$$\mathcal{I}(\mathbf{x}) = \begin{bmatrix} \mathcal{I}(\mathbf{p}) & \cdot \\ \cdot & \mathcal{I}(\mathbf{v}) \end{bmatrix} = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \begin{bmatrix} \mathcal{I}_i(\mathbf{p}) & \cdot \\ \cdot & \mathcal{I}_i(\mathbf{v}) \end{bmatrix} \tag{30}$$

We use $\mathcal{I}(\mathbf{x})$, $\mathcal{I}(\mathbf{p})$ and $\mathcal{I}(\mathbf{v})$ to denote the Fisher information matrix defined by considering only the parameters $\mathbf{x}$, $\mathbf{p}$ and $\mathbf{v}$ respectively. Both $\mathcal{I}(\mathbf{p})$ and $\mathcal{I}(\mathbf{v})$ turn out to be principal sub-matrices of $\mathcal{I}(\mathbf{x})$. In all cases, independent measurements from additional sensors in general positions will never decrease the total information in each $\mathbf{x}$, $\mathbf{p}$ and $\mathbf{v}$.

Proposition 6. *The condition $n \geq 4$ is a necessary condition for $\mathcal{I}(\mathbf{x})$ to be non-singular.*

Proof: Recall that $\mathcal{I}(\mathbf{x}) = \nabla F^\top \mathbf{R_f}^{-1} \nabla F$ or given $\mathbf{R_f} = \operatorname{diag}\left(\sigma_1^2, \ldots, \sigma_n^2\right)$ we have [20]

$$\mathcal{I}(\mathbf{x}) = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \nabla f_i^\top \nabla f_i \tag{31}$$

which is a sum of matrices each with rank at most 1. Now a well-known result states that a rank-k matrix can be written as the sum of k rank-1 matrices but not fewer. This immediately implies our result and completes the proof. ■

Proposition 7. *The following statements concerning efficient estimators hold.*

1) *If n is finite then no efficient estimator exists for $\mathbf{x}$.*
2) *If $\mathbf{p}$ is known and n is finite then an efficient estimator for $\mathbf{v}$ exists and is given by the standard linear maximum likelihood estimator.*
3) *If $\mathbf{v}$ is known and n is finite then no efficient estimator exists for $\mathbf{p}$.*

Proof: This result follows from a general result concerning efficient estimators given in Theorem 1 and 2 in [18]. ■

We do not consider the design of unbiased (but inefficient) estimators for either $\mathbf{p}$ or $\mathbf{x}$ in this work. However, we know that no unbiased estimator for $\mathbf{x}$ exists with a finite variance when $n < 4$. Similarly, no unbiased estimator for $\mathbf{p}$ or $\mathbf{v}$ exists with a finite variance when $n < 2$.

B. Discussion on the Cramer-Rao Bound

The Cramer-Rao inequality assumes an unbiased estimation algorithm and an estimator which achieves the inequality is called an efficient estimator. An efficient estimator does not exist for $\mathbf{x}$ or $\mathbf{p}$ when n is finite (For example, the well-known maximum likelihood localization techniques are only unbiased and efficient when the number of sensors approaches infinity.) but does exist for $\mathbf{v}$ when $\mathbf{p}$ is given. Even if an efficient estimator does not exist then it may be possible to design an unbiased estimator. This possibility is not explored in this work. In practice, a system designer may be constrained in their choice of parameter estimator. Likely, the estimation technique used in practice will be biased [21], [22]. The Cramer-Rao bound for unbiased estimators is still an interesting benchmark with which intuitively pleasing results and performance measures can be derived. However, these results can only be considered as a guide.

It is interesting that the variance (or mean-square-error) of an estimate can sometimes be made smaller at the expense of increasing the bias [23]. The work of [24], [25] explores the concept of bias-variance trade offs in estimation. In [17], [25] a *biased* Cramer-Rao inequality and in [24], [25] a *uniform* Cramer-Rao inequality are developed and can be used to study this so-called bias-variance trade off. These ideas are yet to be fully explored in the localization and target tracking literature.

VI. On the Fisher Information for Velocity Estimation

Doppler-based measurements are often used to estimate the target velocity alone. In this section we explore the relationship between the transmitter and sensor positions and the velocity estimation error lower bound defined by $\mathcal{I}_i(\mathbf{v})$. To this end, consider the sub-matrix

$$\mathcal{I}(\mathbf{v}) = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \begin{bmatrix} \cos^2(\phi_i) & \frac{\sin(2\phi_i)}{2} \\ \frac{\sin(2\phi_i)}{2} & \sin^2(\phi_i) \end{bmatrix} \tag{32}$$

where $\cos\phi_i = \psi_{i,1}$, and $\sin\phi_i = \psi_{i,2}$.

Firstly, we note that optimal sensor placement for velocity estimation is equivalent to the optimal sensor placement for range-based localization as outlined in [26]. However, importantly, the optimal sensor placement for estimating the target velocity does not depend on the velocity itself. Hence, such placements can be used in practice to estimate the velocity of the target when other measures are available for positioning; e.g. we have already discussed a number of hybrid scenarios in which the additional measurements provide access to positioning information as opposed to velocity.

Moreover, with regards to the optimal sensor placement for velocity estimation we assume the error variance in doppler measurements is independent of their true value. While this is not unreasonable in many applications, this assumption can be relaxed such that the variance is different across sensors. Moreover, we can even consider cases in which the standard deviation is multiplied by a percentage of the true range value between the target and individual sensors. In both these cases the Fisher information matrix for velocity estimation becomes very similar to the Fisher information matrix for bearing-only localization and we point to [26] for the details.

We summarize the result on the optimal sensor placement for estimating the velocity of the target in the following proposition.

Proposition 8. *Let the angle subtended at the target by two sensors i and j be denoted by $\vartheta_{ij} = \vartheta_{ji}$. One set of optimal sensor placement is characterized by*

$$\vartheta_{ij} = \vartheta_{ji} = \frac{2}{n}\pi \tag{33}$$

for all adjacent sensor pairs $i, j \in \{1, \ldots, n \geq 4\}$ with $|j-i| = 1$ or $|j-i| = n-1$, and then by a possible application of the following actions on (33):

1) *Changing the true individual sensor-target ranges, i.e. moving a sensor from $\mathbf{s}_i$ to $\mathbf{p} + k(\mathbf{s}_i - \mathbf{p})$ for some $k > 0$.*
2) *Reflecting a sensor about the emitter position, i.e. moving a sensor from $\mathbf{s}_i$ to $2\mathbf{p} - \mathbf{s}_i$.*

For example, Fig. 2 illustrates two optimal sensor placement scenarios obtained from each other by reflecting a particular

sensor about the emitter position.

Insert Figure 2a here

Insert Figure 2b here

Fig. 2. This figure illustrates two scenarios obtained from each other by reflecting a particular sensor about the emitter position. This reflection does not affect the optimality of the sensor-target configuration.

For more information on the proof of Proposition 8 and further details the reader may refer to [26].

VII. Illustrative Example

In this section we demonstrate the performance of the algorithm introduced in Section IV to estimate the position and the velocity of the target. We consider the setting depicted in Fig, 3. We consider the case where the measurements are corrupted by ten different levels of noise, where the noise is assumed to be Gaussian with zero mean and variable variance (and independent at each sensor). The carrier frequency is assumed to be constant for all sensors and equal to 150 MHz. Moreover, the nominal doppler sifts are: $\delta_1 = -9.2664$, $\delta_2 = 3.1640$, $\delta_3 = 3.1640$, $\delta_4 = -7.8162$, $\delta_5 = -9.4941$, $\delta_6 = 7.3782$, and $\delta_7 = 9.1797$ all in Hz. The error in the estimates of position and the velocity for different noise levels after repeating the scenario for one hundred times when six and seven measurements are used are depicted Fig. 4 respectively (In the six measurement case, it is s_7 measurements which are omitted.). Furthermore, we employ a maximum likelihood estimator that is initialized with the solution of the optimization problem (25) to estimate the position and the velocity of the target. The obtained maximum liklihood estimates are presented in Fig. 4. Note, that in the simulations considered, when the estimator is initialized at a random value the estimate does not converge to a value close to the solution for any of the considered noise levels. One notices that using seven measurements the maximum liklihood solution when initalized at the solution of optimization problem (25) outperforms the Cramer-Rao lower bound. We conjecture that it is due to the possibility of obtaining a biased estimate using the methods outlined in this paper.

Additionally, note that the setting presented in Fig. 3 shows elements of a degenerate geometry as well: sensors at positions $\mathbf{s}_2$ and $\mathbf{s}_3$, and the target are collinear. However, here due to the presence of other sensors at non-collinear positions the estimation can be carried out effectively. Due to this issue, if one considers only five measurements taken by $s_1, \ldots, s_5$ the localization problem cannot be solved.

Insert Figure 3 here

Fig. 3. The setting considered in the illustrative example. The squares denote the position of the sensors and the diamond is the target.

VIII. Concluding Remarks and Future Directions

The minimum number of doppler shift measurements necessary to have a finite number of solutions for the unknown target position and velocity is calculated analytically via algebraic arguments. Additionally, we stated the necessary and sufficient

Insert Figure 4a here

Insert Figure 4b here

Fig. 4. The error in the estimate of the position and the velocity of the target using the proposed polynomial optimization method (Polynomial) and maximum likelihood initialized at the solution of the optimization problem (ML-Polynomial) after repeating the scenario 100 times with variable noise variances when six and seven sensors are used.

number of generic measurements to have a unique solution for the target parameters. Later, the same problem has been studied where in addition to doppler shift measurements, other types of measurements are available, e.g. bearing or distance to the target from each of the sensors. Finally, a method based on polynomial optimization is introduced to calculate an estimate for the position and the velocity of the target using noisy doppler shift measurements. A numerical example is presented to demonstrate the performance of this algorithm.

Moreover, some remarks concerning the Cramer-Rao inequality and its relationship to the estimation problem were given. For the case of doppler-based target velocity estimation we completely characterized the sensor-target geometry and provided a number of conditions on the optimal placement of the sensors and the transmitters.

A possible future research direction is to design a dynamical estimator to estimate the position and the velocity of the target measuring the doppler shift measurements continuously. Alternatively, one could consider the notion of constraint-based optimization for localization as discussed in, e.g., [19], [27]–[31].

REFERENCES

[1] L.R. Malling, "Radio doppler effect for aircraft speed measurements," *Proceedings of the IRE*, vol. 35, no. 11, pp. 1357–1360, November 1947.

[2] E.J. Barlow, "Doppler radar," *Proceedings of the IRE*, vol. 37, no. 4, pp. 340–355, April 1949.

[3] Walter R. Fried, "Principles and performance analysis of doppler navigation systems," *IRE Transactions on Aeronautical and Navigational Electronics*, vol. 4, no. 4, pp. 176–196, December 1957.

[4] S.N. Salinger and J.J. Brandstatter, "Application of recursive estimation and kalman filtering to doppler tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 19, no. 4, pp. 585–592, July 1970.

[5] D.C. Torney, "Localization and Observability of Aircraft via Doppler Shifts," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 43, no. 3, pp. 1163–1168, 2007.

[6] Y.T. Chan and F.L. Jardine, "Target localization and tracking from Doppler-shift measurements," *Oceanic Engineering, IEEE Journal of*, vol. 15, no. 3, pp. 251–257, 1990.

[7] M. Amin, P. Zemany, P. Setlur, and F. Ahmad, "Moving target localization for indoor imaging using dual frequency CW radars," in *Sensor Array and Multichannel Processing, 2006. Fourth IEEE Workshop on*. IEEE, 2008, pp. 367–371.

[8] Y.C. Xiao, P. Wei, and T. Yuan, "Observability and Performance Analysis of Bi/Multi-Static Doppler-Only Radar," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 46, no. 4, pp. 1654–1667, 2010.

[9] N. Alam, A.T. Balaie, and A.G. Dempster, "Dynamic path loss exponent estimation in a vehicular network using doppler effect and received signal strength," in *Proceedings of IEEE 71st Vehicular Technology Conference*, 2010, pp. 1–5.

[10] T. Johnsen and K. E. Olsen, "Bi- and multistatic radar," in *Advanced Radar Signal and Data Processing*, 2006, pp. 4.1– 4.34.

[11] B. Kusý, I. Amundson, J. Sallai, P. Völgyesi, A. Lédeczi, and X. Koutsoukos, "Rf doppler shift-based mobile sensor tracking and navigation," *ACM Trans. Sen. Netw.*, vol. 7, pp. 1:1–1:32, August 2010.

[12] D.A. Cox, J.B. Little, and D. O'Shea, *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*, vol. 10, Springer Verlag, 2007.

[13] I. Shames, B. Fidan, B. D. O. Anderson, and H. Hmam, "Cooperative self-localization of mobile agents moving in planar formations," *IEEE Transactions on Electronics and Aerospace*, vol. 47, no. 3, pp. 1926 – 1947, 2011.

[14] I. Shames, P. T. Bibalan, B. Fidan, and B. D. O. Anderson, "Polynomial methods in noisy network localization," in *17th Mediterranean Conference on Control and Automation*. IEEE, 2009, pp. 1307–1312.

[15] D. Henrion and J. Lasserre, "Detecting global optimality and extracting solutions in gloptipoly," in *In Positive Polynomials in Control*, D. Henrion and A. Garulli, Eds., Lecture Notes on Control and Information Sciences. Springer-Verlag, 2005.

[16] P. Parrilo, "Semidefinite programming relaxations for semi-algebraic problems," *Mathematical Programming*, vol. 96, no. 2, pp. 293–320, 2003.

[17] H.L. Van Trees, *Detection, Estimation and Modulation Theory*, John Wiley and Sons, Inc., New York, NY, 1968.

[18] A. Host-Madsen, "On the existence of efficient estimators," *IEEE Transactions on Signal Processing*, vol. 48, no. 11, pp. 3028–3031, November 2000.

[19] A.N. Bishop, B. Fidan, K. Dogancay, B.D.O. Anderson, and P.N. Pathirana, "Exploiting geometry for improved hybrid AOA/TDOA based localization," *Signal Processing*, vol. 88, no. 7, pp. 1775–1791, July 2008.

[20] A. Swami, "Cramer-rao bounds for deterministic signals in additive and multiplicative noise," *Transactions on Signal Processing*, pp. 231–244, September 1996.

[21] V.J. Aidala and S.C. Nardone, "Biased estimation properties of the pseudolinear tracking filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 18, no. 4, pp. 432–441, 1982.

[22] M. Gavish and AJ Weiss, "Performance analysis of bearing-only target location algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 3, pp. 817–828, 1992.

[23] Y.C. Eldar, "Uniformly improving the Cramer-Rao bound and maximum-likelihood estimation," *IEEE Transactions on Signal Processing*, vol. 54, no. 8, pp. 2943–2956, 2006.

[24] A.O. Hero, J.A. Fessler, and U. Usman, "Exploring estimator bias-variance tradeoffs using the uniform CR bound," *IEEE Transactions on Signal Processing*, vol. 44, no. 8, pp. 2026–2041, August 1996.

[25] Y.C. Eldar, "Minimum variance in biased estimation: Bounds and asymptotically optimal estimators," *IEEE Transactions on Signal Processing*, vol. 52, no. 7, pp. 1915–1930, July 2004.

[26] A.N. Bishop, B. Fidan, B.D.O. Anderson, K. Dogancay, and P.N. Pathirana, "Optimality analysis of sensor-target localization geometries," *Automatica*, vol. 46, no. 3, pp. 479–492, March 2010.

[27] R.I. Hartley and P. Sturm, "Triangulation," *Computer Vision and Image Understanding*, vol. 68, no. 2, pp. 146–157, November 1997.

[28] M. Cao, B.D.O. Anderson, and A.S. Morse, "Localization with imprecise distance information in sensor networks," *Systems and Control Letters*, vol. 55, no. 11, pp. 887–893, November 2006.

[29] A.N. Bishop, B. Fidan, B.D.O. Anderson, K. Dogancay, and P.N. Pathirana, "Optimal range-difference-based localization considering geometrical constraints," *IEEE Journal of Oceanic Engineering*, vol. 33, no. 3, pp. 289–301, July 2009.

[30] A.N. Bishop, B. Fidan, B.D.O. Anderson, P.N. Pathirana, and G. Mao, "Bearing-only localization using geometrically constrained localization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 1, pp. 308–320, January 2009.

[31] A.N. Bishop, "A tutorial on constraints for positioning on the plane," in *Proceedings of the 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC'10)*, Istanbul, Turkey, 2010, pp. 1689–1694.